CULTIVATION, ISOLATION AND CHARACTERISATION OF THERMOPHILIC AND ACIDOPHILIC RED ALGAE, CYANIDIALES

A thesis submitted in partial fulfilment of the requirements for the degree of Master's in Science in Biotechnology at the University of Canterbury by Adam Colsell

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Abstract

Cyanidiales is the oldest extant order of the red alga, which diverged from the ancestral lineage of Rhodaphyta approximately 1.4 billion years ago. These asexual, unicellular microalgae are thermophilic and acidophilic, making them unique among all known photosynthetic micro-organisms. These species grow in geothermal streams and soils, at acidities below pH 3.0 and at temperatures exceeding 45°C. This unique physiology means that Cyanidiales have the potential for a variety of biotechnological applications, from pigment production to bioremediation of wastewater. Cyanidiales are found at geothermal sites globally, including New Zealand. However, the diversity and taxonomy of the New Zealand Cyanidiales population has been understudied. *In this MSc thesis, Cyanidiales were sampled and isolated from New Zealand geothermal sites in the Taupō Volcanic Zone. Once isolated genetic, morphological and physiological data was gathered in order to describe the strains and resolve taxonomy.* The physiological characterisation involved determining the temperature range and the pH range of the isolates, as well as carbon and nitrogen source utilization of these isolates. The chemotaxonomic characterisation involved determining the fatty acid profile and polar lipid profile of the isolates. Additionally, to resolve the taxonomy of Cyanidiales, five genes will be sequenced which are four conserved plastid genes; *psaA* (photosystem I P700 chlorophyll *a* apoprotein A1), *psbA* (photosystem II reaction center protein D1), *rbc*L (RUBISCO Large subunit), 16S rRNA gene, and one nuclear gene, the 18S rRNA gene. In total, 10 Cyanidiales representatives were isolated and, with the combination of phenotypic and genetic data, it was determined that the isolates represent three lineages. These three lineages included One *Galdieria maxima* lineage from Russia, as well as two *Galdieria sulphuraria* lineages from a European lineage and an American lineage. The European *G. sulphuraria* lineage had not been observed before in New Zealand, which as such is significant as it indicates that the New Zealand Cyanidiales population is more diverse than previous evidence suggested, despite the geographic isolate of New Zealand. This is strong evidence for multiple Cyanidiales colonization events to NZ and a long-distance dispersal mechanism that is yet to be described. The NZ *G. maxima* strain is also significant as genetic data indicates the closest described relative is *G. maxima* IPPAS P507, however the morphological, physiological and biochemical characteristics of the strain are closer to *Cyanidium caldarium* and *Cyanidioschyzon merolae*. Therefore, these strains are an evolutionary link and an opportunity to revise the nomenclature and the taxonomy. Additionally, this MSc thesis is part of a research program into the development of a proposed photobioreactor that will use New Zealand native Cyanidiales and extremophilic methanotrophs strains in consortium, to sequester the carbon emissions from flue gas. In a desktop assessment, using data from both the literature and this thesis, it was determined that the NZ *G.maxima strain* is the most suitable for application in the proposed bioreactor.

In summary, the data from this MSc thesis contributes to understanding of the dispersal, diversity and evolution of Cyanidiales, both in New Zealand and globally. The characterised NZ Cyanidiales are also an asset to the research and development of New Zealand biotechnologies.

Table of Contents

Table of Figures

Glossary

Acronym or abbreviation	Meaning
ACUF	Algal Culture Collection of University Federico II of Naples
BOD	biochemical oxygen demand
CCMEE	Culture Collection of Microorganisms from Extreme Environments
C-PC	C-phycocyanin
DGMG	Digalactosyl-diacylglycerol
EDTA	Ethylenediaminetetraacetic acid
EST	Expressed Sequences Tag
FID	Flame ionization detector
FAME	Fatty acid methyl esters
GC-MS	Gas chromatography–mass spectrometry
HGT	Horizontal gene transfer
IPPAS	Institute of Plant Physiology, Russian Academy of Sciences
ktCO2-e	kilotons of CO_2 equivalents
LC-MS	Liquid chromatography–mass spectrometry
MAll	Modified Allen's medium
MAll-Ht	Modified Allen's Heterotrophic medium
MGDG	Monogalactosyl-diacylglycerol
ML	Maximum likelihood
NZ	New Zealand
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PG	Phosphatidyl glycerol
PI	Phosphatidylinositol
ppm	Parts per million
***psa*A**	photosystem I P700 chlorophyll a apoprotein A1
***psb*A**	photosystem II reaction center protein D1
***rbc*L**	ribulose-1,5-bisphosphate carboxylase/oxygenase Large subunit
SAG	Culture Collection of Algae at Göttingen University
SQDG	Sulfoquinovosyl diacylglycerol
TLC	Thin-layer chromatography
TVZ	Taupō Volcanic Zone
YNP	Yellowstone National Park

Chapter 1. Background

Global carbon emissions are on the rise, with atmospheric CO_2 concentrations increasing from 280 ppm in the 1700's to greater than 350 ppm in 2005[1,2]. The primary source of these emissions are anthropogenic activities, such as generating energy and transportation[3,4]. In New Zealand, the gross greenhouse gas emissions increased to 80,900 kilotons of CO_2 equivalents (ktCO_2-e) in 2017, which was an increase of 23% since the 1990 value. The primary contributors to NZ emissions are the Energy and Agricultural sectors [4]. The increase in atmospheric CO_2 is the primary driver of Climate Change and ocean acidification [5–7], both of which will have negative effects on the environment, human health, industry and crop security in the near future. As such, there has been a significant effort to reduce emissions where economically viable. One of the potential solutions to reducing CO_2 emissions is the mass cultivation of microalgae, often in place of traditional land plant crops [8,9]. Microalgae, like most land plants, are photosynthetic and hence reduce atmospheric CO_2 into sugars, carbohydrates and/or lipids using the energy from the sun. There are several advantages to using microalgae over traditional crop plants. Compared to land plants, microalgae have higher growth rates, greater photosynthetic efficiency and a higher proportion of the biomass is digestible. As such, more edible mass can be produced in the equivalent space in a shorter time. Mass microalgae cultures can yield up to 20 kg/m^2 /year [10] which equates to about 200,000 kg per hectare. For comparison the yields for many cereal crops, like maize, equate to around 10,000 kg per hectare [11]. Additionally, microalgae cells can have a greater percentage of carbohydrates or lipids per gram of dry weight than land plants, up to 70% depending on the species and cultivation conditions [12,13]. For context, the combination of these factors means that a microalgae with 50% lipid by dry weight would have a theoretical oil yield of around 100,000 litres of oil per hectare, whereas Palm oil yields are around 5500 litres of oil per hectare [13]. Microalgae also produce novel high-value products and have tolerance to high CO_2 concentrations, making them useful in bioreactors [14,15]. In the early 2000's, the global commercial microalgal biomass market produced around 5000 tonnes of dry mass per year, which was worth an estimated 1.25 billion US dollars before processing [15,16]. Most common commercial strains are species of *Chlorella*, *Dunaliella*, *Arthrospira* and *Haematococcus* [13,15]. These species are used to generate a wide range of products, from human food components, pigments such as carotenoids and astaxanthin, and livestock feed[9,13–16]. The feed industry, by biomass generated, is the largest sector of the commercial microalgae industry, accounting for around 30 percent of global harvest [17].

While carbon dioxide is considered the principal greenhouse gas, methane (CH_4) has a global warming potential (GWP) 34 times more potent than CO_2 over a hundred years with climate-carbon feedback [18](Note here that most governments, including New Zealand's, use the 2007 standard (Kyoto Protocol) where methane has a GWP of 25 times CO_2 [4,19]. Methane is a major component of New Zealand's gross greenhouse emissions, with about 42% as methane and 45% as CO_2. Most emission sources generate low concentrations across a wide area e.g. cattle farms or transport, which are difficult to control, manage and/or capture. There are also a number of emission sources in NZ that are highly concentrated and generated from a point source. Examples of these sources are oil and gas refineries, which in NZ vent approximately 184 million cubic metres of waste gas, and geothermal power plants which produced 814 kt CO_2-e in 2017 [4]. These flue gases, generated from a point source, can be easily captured and have the potential to be used as a carbon source for organisms such as microalgae. Growth of microalgae with flue gases will sequester the carbon emissions, and depending on the organism, could be used to generate a valuable product such as feedstock or biofuel.

1.2. MSc Research Project context

A group of CAPE/Scion researchers have recently received funding from the Ministry of Business, Innovation and Employment (MBIE) through an Endeavour Fund Smart Idea grant for a proposal [20] that uses microalgae to reduce carbon emissions (CO_2 and CH_4) from the geothermal and fossil fuel industries. Their proposal was to develop a bioreactor that uses a *Cyanidiales* microalgae strain as a photoautotroph, and a Verrucomicrobia bacteria strain as a methanotroph, to completely convert the carbon emissions from a concentrated point source, such as a geothermal power plant, into biomass. The biomass could then be processed into a profitable product such as biofeedstocks, especially for the dairy industry. The basic principle behind the proposed research programme was that photoautotrophs will use CO_2 from the waste emissions and reduce it using sunlight energy to generate biomass and O_2. The methanotroph will use O_2, supplemented from the photoautotroph, to oxidize the methane and other short-chain hydrocarbons for biomass and energy generation [20]. This not only produces additional biomass but also more CO_2 for the photoautotroph, hence this sets up a cyclic bioreactor where the inputs are simply waste gas emissions and natural sunlight, and the output is biomass, ideally as carbohydrates and proteins (see figure 1.1). The biomass could then be processed into a commercial product such as feedstock for livestock. This project has the potential to reduce greenhouse gas emissions, while generating a profitable product out of a currently zero value waste by-product.

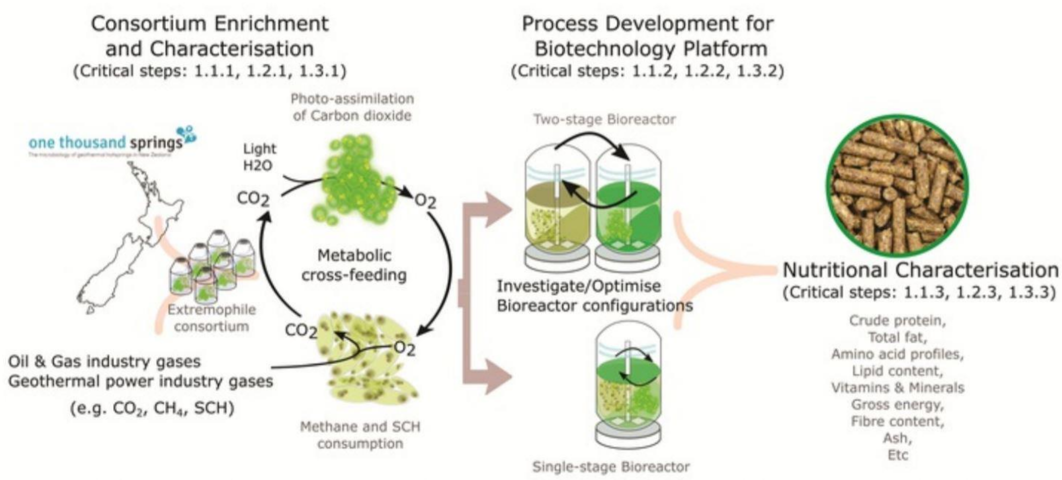


Figure 1.1: *: Diagram of the proposed research path of the MBIE project, with the first step involving isolation of the desired microorganisms from NZ, then candidate isolates will be used to develop and optimise the bioreactor, for growth and nutritional value. Sourced from Carere C.R. (2016).*

Selection of organisms and strains is a key step in the development of a bioreactor. In the MBIE Smart Idea project, the researchers have chosen to use extremophilic photoautotrophs and methanotrophs. This affords the process a number of advantages over standard bioreactor setups. For example, the researchers are using thermophilic acidophiles whose optimal growth conditions are temperatures above 45°C and pHs below 3. Operating under these conditions greatly diminishes growth of undesired strains [21] that would decrease the efficiency of the bioreactor, thus affecting profit and the economic viability of the bioreactor. Additionally, the researcher will be using non-genetically modified and New Zealand native strains, as this avoids the need for regulatory approval as required for biosecurity within New Zealand. The selection of compatible phototrophic and methanotrophic strains is critical for the success of the project. Up until recently, all bacterial methanotrophs were mesophilic strains from the Proteobacteria phylum, so growth temperature ranges were limited to approximately 5-40°C and pH growth ranges were limited to approximately 4.5 to 7.5 [22]. However, recently there has been the discovery of a number of methanotrophic strains from the phylum Verrucomicrobia (genera *Methylacidiphilum* and *Methylacidimicrobium*) that are extremophilic, and several have been isolated from New Zealand. The strain *Methylacidiphilum* V4, for example, has an optimal growth temperature of 60°C and a pH optimum of around 2-2.5 [22,23]. Hence these strains are the prime candidates for the role of methanotroph within the required bioreactor operating parameters. The selection of the photoautotroph is also challenging, as species of phototrophic prokaryote microorganisms that can survive under pH 5 are limited [33–35], and at this time, there are no known examples that can grow under pH of 4 [33–35]. This limits the selection to strains of the microalgae order Cyanidiales, as these are the only known photosynthetic microorganisms that grow at a combination of high temperatures and low pH. Cyanidiales have previously been observed in New Zealand [24–27]. However, access to these strains is uncertain as it is unclear whether permission (DOC, Māori or otherwise) to collect samples was originally obtained and/or whether there is commercial freedom to operate. While these strains can be used for early-stage trials and development, additional NZ Cyanidiales isolates will be needed if the proposal moves to a commercial or later stage of development. Thus, the researchers need additional Cyanidiales isolates, and the greater the number and diversity, the higher the probability of this MBIE project moving into later developmental stages. Broadly, my MSc research project centres around the need to enrich, isolate and characterise new strains of Cyanidiales.

My role will be to;

1. Sample geothermal sites in New Zealand for Cyanidiales.
2. Enrich for and isolate Cyanidiales strains.
3. Carry out basic characterisation, and phylogenetic analysis to determine taxonomy.
4. Assess the individual isolates suitability for use in the theoretical bioreactor.

Chapter 2. Cyanidiales- A Literature Review

2.1 Introduction

Cyanidiales is the oldest extant order of Rhodophyta (Red algae), diverging from the main lineage approximately 1.1 to 1.5 billion years ago [28,29]. All members are small (~10µm), asexual unicellular autotrophs and produce the photosynthetic pigments chlorophyll A and c-phycocyanin [30,31]. Therefore, under autotrophic conditions, these algae appear blue-green rather than red. Many members of this order are obligate acidophilic thermophiles, requiring acidic conditions with pH below 4 and elevated temperatures between 40°C to 56°C[31]. As such, Cyanidiales thrive in acidic sulphuric geothermal sites globally [26,32]. This polyextremophilism makes this order notable, as no other known photosynthetic microorganism can survive this combination of conditions[33–35]. Due to living in the harsh environments of geothermal sites, many have adaptations to cope with other stress, such as heavy metals, high salt (up to 10% NaCl) [31,36] and atmosphere saturated with pure CO_2 [37].

Taxonomy of the Cyanidiales historically has been challenging due to the Cyanidiales simple morphology, microscopic size and blue-green appearance, meaning species has been erroneously characterised in other photosynthetic phyla as Cyanobacteria and as Green algae. Currently, there are only three accepted genera in this order, *Cyanidium, Cyanidioschyzon* and *Galdieria*, which are primarily characterised by simple morphological, biochemical and ecophysiological features [30,31]. However, there is molecular evidence to suggest that there is far more diversity in the Cyanidiales than the current taxonomic system suggests [38], and that certain strains have been classified erroneously[29,38]. Despite this molecular evidence, family and genera descriptions have received only minor amendments and revisions since the original description[30]. This is reviewed more in section 2.2.

Research on Cyanidiales, both historically and contemporary, has been multi-disciplinary and has contributed to fields outside of just phytology or microbiology. Cyanidiales have been key in both historical and contemporary theories on the origin and evolution of plastids, Rhodophyta and Archaeplastida(photosynthetic eukaryotes) in general [28,29,39]. Cyanidiales, *Cyanidioschyzon* in particular, are considered among the simplest photosynthetic eukaryotes and are model organisms for understanding the replication of plastids and mitochondria [40,41]. Additionally, *Cyanidioschyzon merolae* 10D was the first complete algal genome to be sequenced [42,43], and later genome sequencing of *Galdieria* strains has provided strong evidence for evolutionary driven horizontal gene transfer in Eukaryotes [44–46].

Cyanidiales strains, as extremophilic algae, are a prime candidate for the development of microalgae biotechnology such as pigment production, bioremediation and wastewater treatment [47]. Discussed in section 2.5

Cyanidiales are globally distributed and have been scientifically documented in New Zealand geothermal sites since the 1970's [24,25], however modern scientific isolation and characterisation of Cyanidiales has been limited to a single research article. This means that the dispersal phylogeny and any potential development of biotechnology using native strains would be limited. This thesis by isolating and describing New Zealand Cyanidiales, contributes not only to the understanding of the dispersal, diversity and evolution of Cyanidiales but also to the future research and development of New Zealand biotechnologies.

2.2 Taxonomy of Cyanidiales- The Historic and current challenges

2.2.1 Historical Taxonomy

The taxonomy of Cyanidiales has presented numerous challenges historically and currently; early phytologists have documented blue-green biomass in sulphuric acid geothermal sites since the 19th century as detailed in secondary sources[30,48–50]. Historically it was difficult to distinguish between microalgae without modern microscopes or biochemical and molecular techniques. As such, these Cyanidiales were often misclassified as either Cyanophyta (blue-green, currently Cyanobacteria) or Chlorophyta (Green algae)[48,50,51]. The Cyanidiales are also distributed globally, meaning different populations of Cyanidiales were being discovered and named independently. Therefore Cyanidiales spp. were often classified under numerous binomials simultaneously[52,53]. The most common binomial was *Cyanidium caldarium*, which was reclassified correctly as a Rhodaphyte, given it was a photosynthetic eukaryote, produced floridean starch and lacked chlorophyll B [52].

These *C. caldarium* cultures were not axenic cultures and two strains were isolated and referred to as *C. caldarium* forma A and forma B. Comparative studies between these two strains, such as the fatty acid profile [54], would determine that they represented different species. This, and the discovery of a novel Cyanidiales species, *Cyanidioschyzon merolae*[55], would lead to the taxonomic revision of *Cyanidium* and the start of modern Cyanidiales nomenclature and taxonomy[30,31,49].

The current nomenclature and taxonomy of the Cyanidiales was described in Merola *et al.* 1981[30] which established the class Cyanidiophyceae and placed it into the phylum Rhodophyta. Merola *et al.* 1981 first split the *Cyanidium* forma A and forma B into two genera, *Cyanidium* and *Galdieria* respectively. Additionally, Merola *et al.* 1981 placed the recently discovered *Cyanidioschyzon merolae* into the class Cyanidiophyceae. This established the current nomenclature with the three genera and the respective species, *Cyanidium caldarium, Galdieria sulphuraria* and *Cyanidioschyzon merolae*. The different genera descriptions were based on differences in observed morphological, ecophysiological and biochemical features. *C. merolae* and *C. caldarium* were both placed in the new family Cyanidiaceae as both were described as strictly autotrophic, with a single mitochondrion, but lacking vacuole or trienoic acids such as α-linolenic acid. However, *C. caldarium* was described to have 2-6μm round cells which reproduce via 4 autospore/endospores, while *C. merolae* cells were described as oblong or club-shaped at 1-1.5μm wide and 3-4μm long and which reproduce via binary fission. It would be determined later that the *C. merolae* lacked a cell wall which is hypothesised to cause the different cell shape and reproductive pattern [56,57]. *G. sulphuraria* was described as a facultative heterotroph with larger round cells (4-11μm) that contained a vacuole, multiple mitochondria, could produce trienoic acids and reproduced with 4-32 autospores . Based on the difference in the descriptions of *Galdieria* compared to the other two species, Merola *et al.* established a separate family Galdieriaceae. Figure 2.1 illustrates the morphological differences between the genera.

Table 2.1: Basic diagnostic features of Different Cyanidiales species[31,36,68]

Feature	C. merolae	C. caldarium	G. sulphuraria
Shape	Oval, club-like	Spherical	Spherical
Size (μm)	1.5x4	2-6	4-11
Reproduction	Binary fission	Endosporse	Endospores
Vacuole	Absent	Absent	Present
Cell wall	Absent	Present	Present
Facultative heterotrophy	No	No	Yes
Nitrogen sources	NH_4^+, NO_3^-	NH_4^+, NO_3^-	NH_4^+
α-linolenic acid (trienoic acids)	Absent	Absent	Present

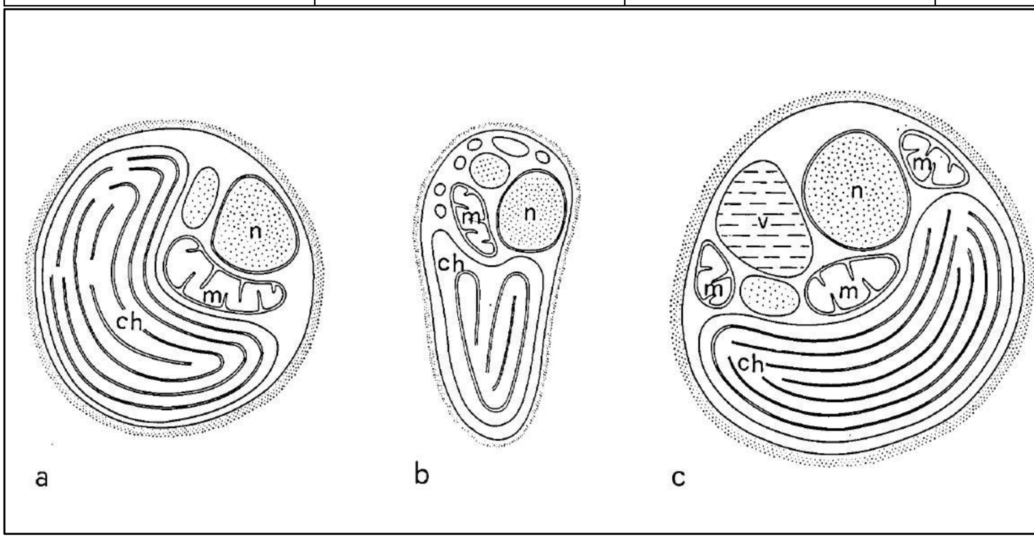


Figure 2.1: Illustration showing the Cell morphology of a) *Cyanidium caldarium* b) *Cyanidioschyzon merolae* and c) *Galdieria sulphuraria*. Basic organelles are shown, n-nucleus, ch-chloroplast/plastid, m-mitochondrion, and v-vacuole. Sourced Merola *et al.* (1981)

There was some disagreement with the new taxonomy [52,58–60], for example Ott and Seckbach 1994, suggested that *Galdieria* should stay as *Cyanidium* [52,58] and that *Cyanidioschyzon* should be placed in the "primitive" unicellular Porphyridiaceae family, as it reproduces asexually by binary fission [52]. Although the new taxonomy was generally accepted and a number of laboratory "*Cyanidium*" strains were reviewed for heterotrophic growth and were determined to be mixed *Galdieria-Cyanidium* cultures[55]. This led to the descriptions of new *Galdieria* species; *G. maxima, G.daedala* and *G.partita* [61]. Additionally, a review on mesophilic and neutrophilic "blue-green algae" strains found in caves in Chile, Italy and France showed the strains matched the descriptions of *Cyanidium* [62,63]. At the time, these strains received no official description at the time and were placed as a sub-species *Cyanidium caldarium* var. chilense and were known colloquially as “Cave *Cyanidium*”. Since the discovery of the "Cave *Cyanidium*", additional examples have been described [64–66] . Phylogenetic placement of these representatives confirms the lineage to be distinct to *C. caldarium* [29,38] and has been given the official name of *Cyanidium chilense* Schwabe [67].

Later attempts to update the taxonomy simply added additional data like nitrogen sources and found that *Cyanidium* and *Galdieria* did have enough significant differences to be classified as different genera, [31,36,68] mostly in agreement with Merola *et al.* 1981 (see table 2.1). However there were concerns that some of the diagnostic character sets were strain or condition dependent [36,59]. For example, while all Cyanidiales strains that could only utilize NH_4^+ as a Nitrogen source were *Galdieria* strains, many other *Galdieri*a strains could utilize nitrate as well[36,59]. Meanwhile the lipid profile of Cyanidiales strains can be altered by culture conditions, such pH [69] and presence of carbon sources [70,71]. This means the over reliance on too few characteristics will lead to misclassification of Cyanidiales strains. Although contemporary diagnostic sets, if used correctly, could easily distinguish between different genera, it could not reliably be used to distinguish between *Galdieria* species [36]. Resolution of Cyanidiales taxonomy to the species level would require molecular and phylogenetic data.

2.2.2 Molecular data and the taxonomy of Cyanidiales

As sequencing technology developed, molecular and phylogenetic data for Cyanidiales became available. Initially, it was simple phylogenetic trees and pairwise distance matrices based on partial single gene sequences, such as Rubisco [36,68] or 18S rRNA gene [72]. The data confirmed that the three genera should remain separate but that most of the individual *Galdieria* strains were indistinguishable from each other, except for *G. maxima* [72]. Advancement in sequence technology allowed phylogeny to be derived from multiple sequences which gave better resolution of not only the Cyanidiales clades, but also phylum and kingdom level clades. Data sets ranged from three plastid protein gene sequences, such as *psa*A (photosystem I P700 chlorophyll a apoprotein A1), *psb*A (photosystem II reaction center protein D1), and *rbc*L (ribulose-1,5-bisphosphate carboxylase/oxygenase Large subunit), [38,39] to nine gene sets that contain both nuclear and plastid coded sequences [29]. Often molecular data from Cyanidiales strains was collected and used when focus of the research was not Cyanidiales phylogeny. Rather the focus of the research varied from understanding the origin of plastids [39,73] to understanding the phylogeny and evolution of Rhodophyta, [29] and photosynthetic eukaryotes in general [28].

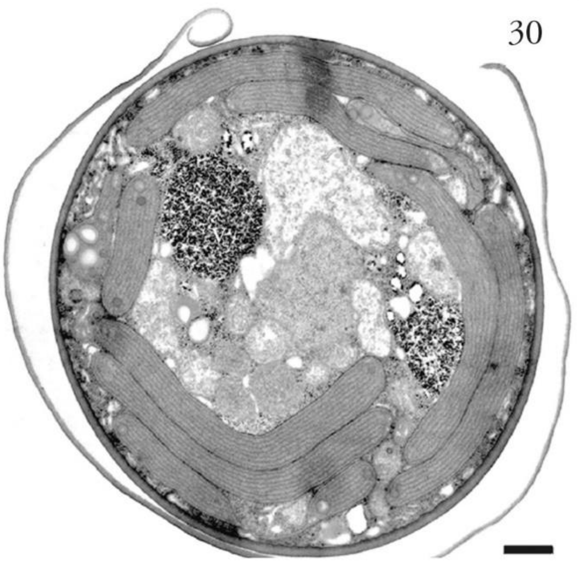


Figure 2.2: TEM microscope image of young *Galdieria maxima* strain P507 cell, scale bar is 1 µm. Sourced from Pinto *et al.* (2003).

In terms of the Cyanidiales phylogeny, this wealth of molecular data confirmed the Cyanidiales lineage was monophyletic and belonged to Rhodophyta, but was first to diverge from the Rhodophyta lineage (1.3-1.5 Bya[28,74]) after the initial Archaeplastida split (1.6-1.8 Bya [28,74]). Additionally, the clades of *Galdieria, Cyanidioschyzon, Cyanidium* and Cave *Cyanidium* were distinct and resolved. The data confirmed separate families of Galdieriaceae and Cyanidiaceae as *Cyanidium* and *Cyanidioschyzon* formed a monophyletic clade separate from *Galdieria* clade. This proved the robustness of the original descriptions for *Cyanidioschyzon, Cyanidium* and *Galdieria* [30], with the exception being strain *G. maxima*. Phylogenetic studies that include molecular data from *G. maxima* [28,29,38,75], found that *G. maxima* would consistently form a sister group with *Cyanidioschyzon* strains. This is significant as *G. maxima* shares many morphological, biochemical and ecophysiological features with *G. sulphuraria*, hence often regarded as a variant of *G. sulphuraria* [36,59,68]. For example, *G. maxima* has a cell wall, larger round cells (10-16µm (fig. 2.2)), has vacuoles, reproduces via autospores and is a facultative heterotroph which differs greatly in comparison the nearest relative *C. merolae* (table 2.1 and fig. 2.1, B) The evolutionary distance between *C. merolae* and *G. sulphuraria* is substantial, as the lineages are estimated to have diverged approximately a billion years ago [28,29,44] and has been likened to the evolutionary distance between *Drosophila* and *Homo sapiens* [44]. This difference in physical features, while being genetically similar, posed a problem to the understanding of the evolution of Cyanidiales. It was theorized *C. merolae* represented the ancestral morphology with primitive features like a lack of a cell wall and vacoule. The close relationship of *C. merolae* and *G. maxima,* and the monophyletic nature of the Cyanidiaceae, suggest the hypothesis that ancestral Cyanidiales morphology would be closest to *C. caldarium,* and that *C. merolae* lost its ability to build a cell wall due to living in an osmotically stable environment and *G. maxima* evolved heterotrophy independently of *G. sulphuraria*. While this relationship remains unresolved, evidence from whole-genome data, discussed in greater detail in section 2.4, suggests that genome reduction in Rhodophyta [76] and horizontal gene transfer are a significant part in the evolution of Cyanidiales [44–46,57]. The description and nomenclature are yet to be amended, and the strain is still referred to as *G. maxima* despite the molecular evidence to suggest it is not a *Galdieria*.

To alleviate the confusion, it was suggested that Cyanidiales should be separated into different lineages for reference until the taxonomy could be resolved. Lineages were based on the phylogeny from sequenced datasets of three concatenated plastid protein genes (*rbc*L, *psa*A and *psb*A)[38]. The four main lineages would be; The true *Galdieria sulphuraria* lineage, The *Cyanidium* lineage, The mesophilic “Cave" *Cyanidium* and The *G. maxima/C. merolae* lineage[32,38](figure 2.3). The true *Galdieria* lineage can be divided further into a non-endolithic lineage *Galdieria*-A (or *Galdieria sulphuraria* lineage) and an endolithic lineage *Galdieria* B (or *Galdieria phlegra* lineage)[38,75]. A number of strains in endolithic lineage have been given the new species description and nomenclature of *Galdieria phlegra* [77], while the *Galdieria* A lineages are likely all variations of or derived from *G. sulphuraria*. This initial distinction is primarily based on molecular data [77], and while it is supported by recent molecular data that includes whole-genome data [45,46], caution is suggested as revision of other related strains is still in progress and it is uncertain if endolithic growth is an accurate diagnostic feature to distinguish between traditional *G. sulphuraria* strains and *G. phlegra* strains. *Galdieria maxima/C. merolae* can be separated into two lineages given the great differences in morphology, biochemical and physiology, and their relevance to fields such as biotechnology. So four to six lineages can be referenced depending on the relevance to the topic.

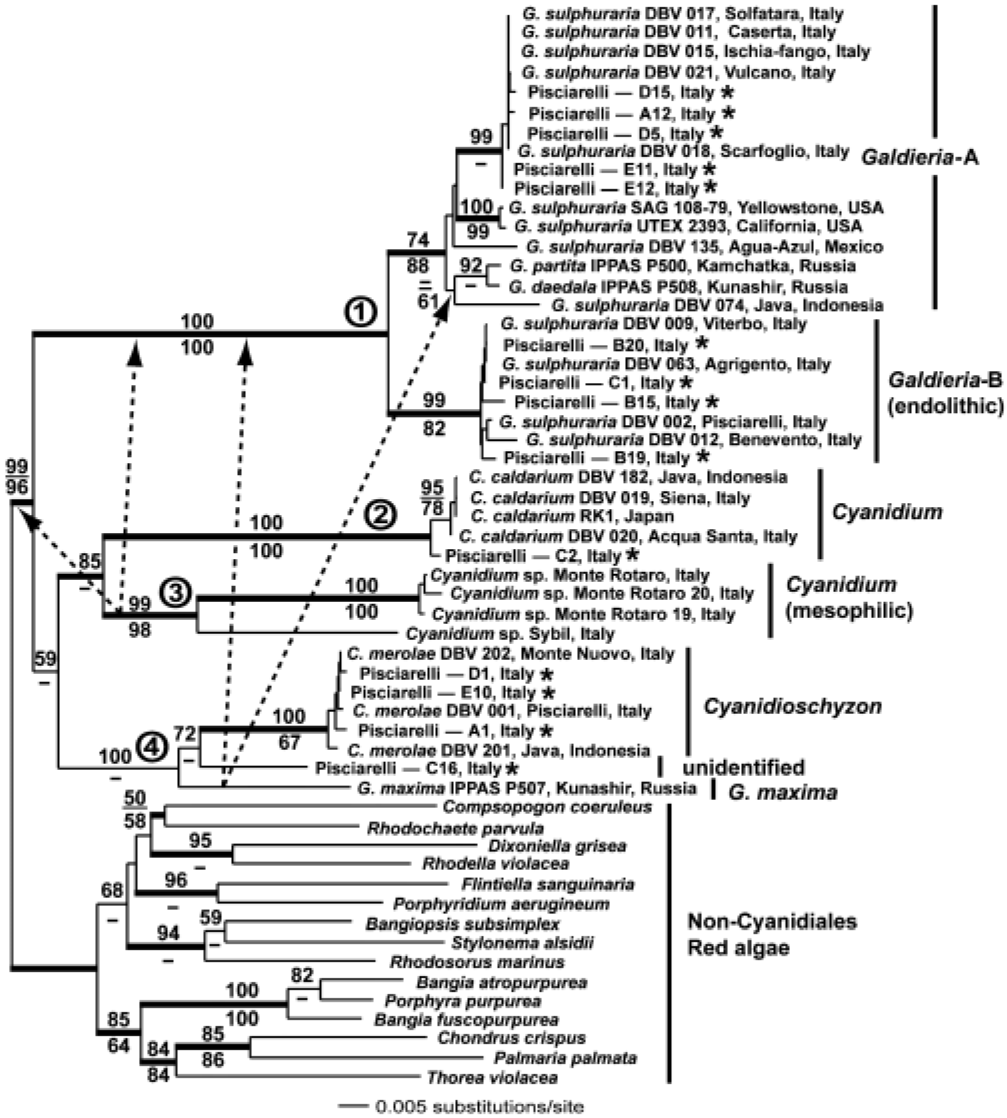


Figure 2.3: Phylogeny of Cyanidiales inferred from minimum evolution(ME) using *rbc*L sequences. Bootstrap values are shown on the branches, bootstrap values above branch is from ME inference of *rbc*L nucleotide sequence and values below the branches are boostrap values from maximum likelihood (ML) inference of the *rbc*L protein sequence. The six Cyanidiales lineages are labelled. Sourced form Ciniglia *et al.* (2004)

As research continues, Cyanidiales are being sampled from new locations such as Taiwan [78]. Many of these new isolates have novel characteristics such as the acidophilic mesophile strains from a burning coal pile, in the Czech Republic [79]. These novel strains will provide insights into the evolution, diversity and phylogeny of Cyanidiales, but will likely be difficult to classify accurately if phenotypes are significantly different to type strains. The accurate classification of novel Cyanidiales strains requires a diverse set of data from morphological, biochemical and ecophysical tests, in combination with accurate molecular data and thorough phylogenetic analysis. The aim of this thesis is to sample, isolate and characterise Cyanidiales from New Zealand to contribute to the ongoing effort to provide clarity to the taxonomy, dispersal and evolution of Cyanidiales in NZ and globally. New Zealand has been undersampled and there is a high probability of isolating novel strains with novel phenotypes (discussed in section 2.3). To ensure accurate classification, a wide range of characteristics need to be determined in conjunction with phylogenetic data on multiple gene sequences.

2.3 Cyanidiales In New Zealand

Research on New Zealand populations of Cyanidiales is severely limited. As such, questions on the origins, evolution and diversity of NZ Cyanidiales remain, which will hamper the development of Cyanidiales in biotechnologies such as waste gas remediation in NZ and understanding of Cyanidiales globally. While Cyanidiales strains have been scientifically documented in New Zealand since at least the 1970's [24,25,80], much of this work has been focused on either microbial ecology or geological studies. Cyanidiales strains were typically identified by microscopic imaging and characterization was either non-existent or extremely limited. To the best of my research, modern published studies that isolated, characterised and produced molecular data from NZ Cyanidiales, is limited to two articles [26,27]. The aim of Donachie *et al.* 2002[27] was to understand the microbial diversity of a single geothermal site, White Island. In the study, a number of diverse microbes were isolated from the hydrothermal water and identified; from Proteobacteria, green-sulfur bacteria, firmicutes and a single Cyanidiales strain (NZ4). From both light and confocal microscope images (fig. 2.4), it was observed that the cells of this strain were round, 2.5-6μm wide, produce 10-15 endospores (described as daughter cells in the article) and cell walls can be observed. It is noted that this strain did grow in a medium with a pH of 0.2. It was confirmed as a Cyanidiales strain due to the pigment profile, and the plastid 16S rRNA sequence data showed a 99.3% similarity to *C. caldarium* [27]. The characterization of this strain was limited to that data and did not include data such as heterotrophic growth or nitrogen sources.

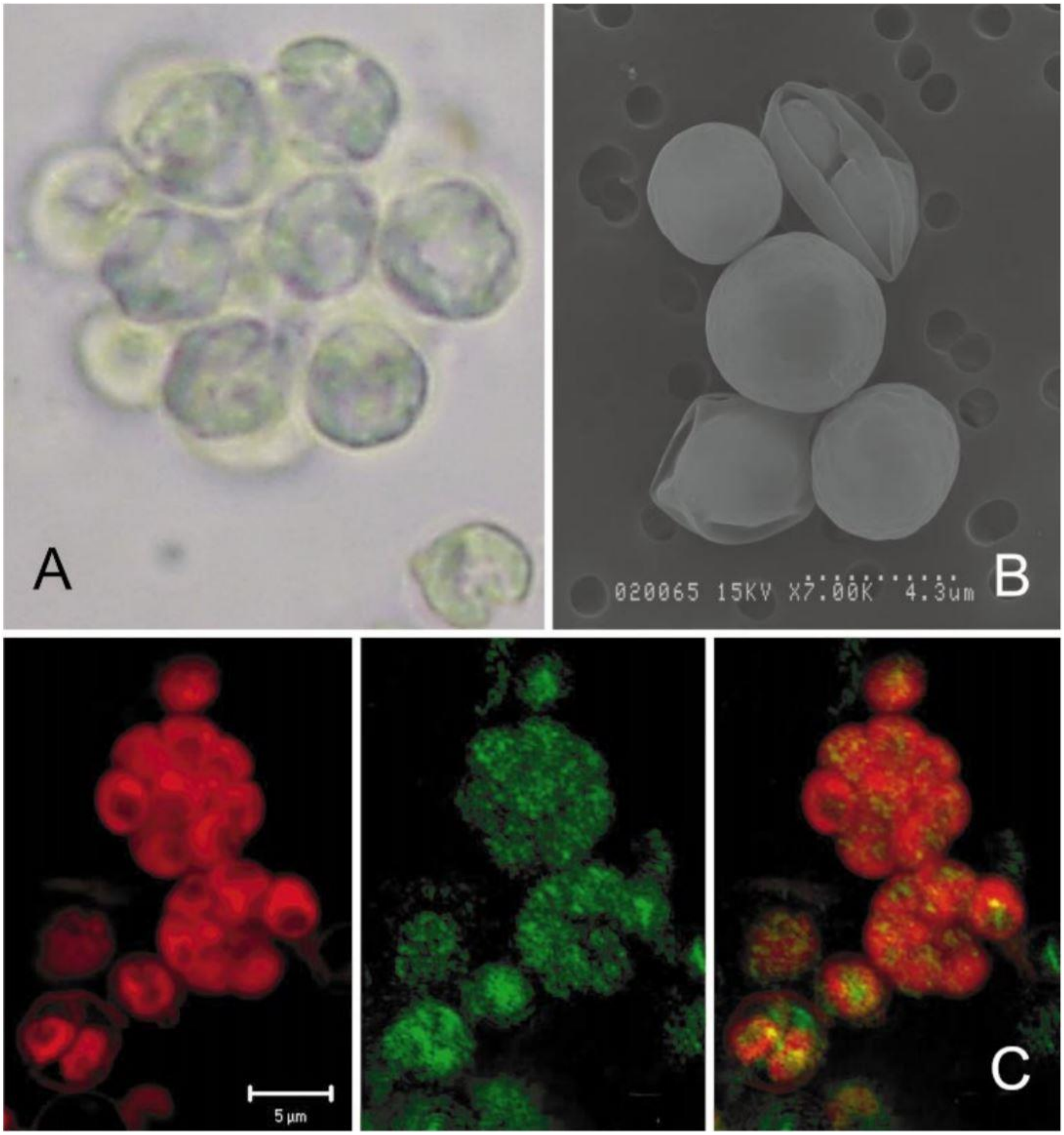


Figure 2.4: Microscope images of New Zealand Cyanidiales strain NZ4. A) Is image from a light microscope B) Is an Image under an Electron microscope C) Are images from Confocal microscope. Endospores are visible in B) and C). Sourced from Donachie *et al.* (2002)

Toplin *et al.* 2008[26] is a more recent study, the focus was on understanding the biogeography and phylogeny of Cyanidiales on a more global scale. Around 60 Cyanidiales strains were isolated from Yellowstone National Park (YNP), Japan and seven sites in the North Island, NZ. Phylogenetic analyses were carried out on partial *rbc*L and 18S rRNA gene sequences. Characterization was limited to observation of morphology under a light microscope for most strains, except for a selection of YNP strains that were qualitatively tested for heterotrophic growth on glucose and nitrogen sources, [26]. The NZ isolates were separated into three type strains based on phylogeny and morphology, Type IV, V and VI. Based on physical morphology from light microscopic imaging (fig. 2.5), NZ Type IV and VI were described as "*G. sulphuraria*-like" with large round cells and appeared similar to YNP Type II isolates, although NZ type VI was slightly smaller (fig. 2.5, C, G & I). NZ type V was described as "*G. maxima*-like " with smaller cells and had a similar appearance to the Japan isolates (fig. 2.5, D-F, H). It was noted that the cell did appear "somewhat oblong". The molecular data was presented in two forms, sequence similarity based on BLAST and phylogenetic trees. The molecular data only partially supports the morphological observations. The *rbc*L tree (fig. 2.6) shows that NZ type IV forms a sister clade with YNP type II and *G. sulphuraria* SAG108.79, meanwhile NZ type V and the Japan isolates all form a clade that include *G. maxima* IPPAS P507. NZ Type VI forms a sister clade with the *G. maxima* clade, contrary to the morphological observations. The data from the *rbc*L tree is supported by the BLAST results of the *rbc*L sequence (table 2.2). NZ type IV closest relative is *G. sulphuraria* SAG108.79 (96%), while NZ type V and VI closest relative is *G. maxima* IPPAS P507, with 93% and 91% respectively. The 18S rRNA gene sequence molecular data conflicts both the morphological observations and the rbcL molecular data, with respect to the NZ isolates. In the 18S rRNA gene tree (fig. 2.7), the "*G. sulphuraria-like*" NZ type IV isolates collapses into the *G. maxima* clade with the Japan isolates and NZ type V isolates. The NZ type VI isolate is not represented on the 18S rRNA gene tree. The nearest BLAST relative based on the 18S rRNA gene sequences is *G. maxima* IPPAS P507 at 99% for all NZ type strains (table 2.2). It is important to note that the NZ type VI is only represented by a single isolate, '5713 Waiotopu', and only the *rbc*L sequence of that strain is published, which is likely why it is absent from the 18S rRNA gene tree. The discrepancy between the molecular data for the NZ isolates, in particular NZ type IV, is noted by Toplin *et al.* 2008. They suggest two possible hypotheses; (1) the discrepancy is due to the difference of evolutionary scale between plastid sequences (*rbc*L) and nuclear sequences (18S rRNA); or (2) that the discrepancy suggests HGT between *G. sulphuraria* relative and a *G. maxima*.

There is strong evidence to support the hypothesis that the evolution of plastid DNA and nuclear DNA do occur at different rates, Toplin *et al.* cited [81,82] as evidence. Additionally, *rbc*L is a protein coding sequence and the third codon positions of Cyanidiales rbcL does show mutational saturation [36]. So typically, to avoid the misleading effects, either residue sequences are used or the evolutionary model is altered (e.g. exclude third codon) [28,38,75]. However, the stated methods suggest Toplin et al did neither. For NZ type V and VI strains, this could explain the lower similarity of *rbc*L to *G. maxima* compared to the 18S rRNA similarity, so the evolutionary distance between *G. maxima* and those strains is less dramatic than the *rbc*L sequence suggests. For the *rbc*L of NZ type IV to be closer to *G. sulphuraria* than *G. maxima*, given the suggested evolutionary distance between them, seems improbable. It would suggest convergent evolution, which if non-random would suggest that *G. sulphuraria rbc*L phenotype differed enough to increase the fitness of NZ type IV in some way or situation. However, the HGT hypothesis seems even more improbable. As stated in section 2.4, genome sequence data provides strong evidence that HGT occurs in Cyanidiales strains. However, the commonly retained HGT are acquired from prokaryotes and provide a clear fitness benefit in geothermal sites, such as detoxifying metals or extracellular transporters that function in acidic environments. The hypothesis suggests that HGT occured between one eukaryote plastid to another eukaryotic plastid with a gene which both parties share, and which would be phenotypically similar. While not impossible, the hypothesis does not stand up to much scrutiny. A simple hypothesis, which the authors would not state for obvious reasons, is the potential of a mixed culture and that there is either prefferation DNA extraction or primer binding. The authors seem aware of these possibilities when discussing the phylogeny of YNP Type IA. The YNP Type IA isolates have "*G. maxima*-like" morphology (fig. 2.5, A), while have 99% identity to *C. merolae* 10D for both *rbc*L and 18S rRNA gene sequence (table 2.2). This is a notable strain as it may provide the link between *G. maxima* and *C. merolae*, helping to resolve the phylogeny. The authors stated concern that sequence similarity could have been from *C. merolae* contamination, so they took a number of precautions such as isolating single cells using a dissecting scope. It is not stated if this level of caution was used in isolation of all strains, although if the same caution was applied to NZ type IV then contamination is unlikely. To test these hypotheses will require isolation of additional NZ type IV and V samples and further characterization and molecular data to confirm classification.

No *C. caldarium* related strains were isolated in the more extensive sampling of Toplins *et al.* 2008 as the earlier plastid 16S rRNA gene data would suggest [27]. While it is very possible that the *Cyanidium* was not sampled or did not enrich, it is also possible that the strain might have been misidentified. A similar situation occurred in YNP. Previous 16S rRNA molecular surveying suggested that the majority of algae in YNP were *Cyanidium* strains, given close similarity (98-99%) to *C. caldarium* chloroplast 16S rRNA [83]. However, the extensive sampling of YNP, including 18S surveys, are yet to indicate the presence of any *C. caldarium* [26,84,85]. Even If no *C. caldarium* strains are isolated as part of this MSc thesis it would be prudent to determine the similarity of the plastid 16S rRNA gene to Cyanidiales in the NCBI database using BLAST search. Given the hypothesis that the 16S rRNA sequence is too conserved to distinguish between genera in Cyanidiales it is possible the strain identified in Donachie *et al.* 2002 is a Galdieria strain and the isolated in this thesis This can be confirmed by sequencing the 16S rRNA genes of the isolates and determining closest relative. Regardless, the isolation of Cyanidiales in New Zealand so far suggests there are between 2 to 3 native lineages. If this were to be the case, then this would has implications for the biogeography and dispersal of Cyanidiales globally as prior to sampling Cyanidiales in NZ, the co-existence of *G. maxima* and *G. sulphuraria* was only documented in Russia, Japan and Iceland [26,32,86]. Based on the *rbc*L phylogeny, [26,32] there is evidence that there has been at least two separate migration events of Cyanidiales to NZ; *G. maxima* from Japan or Russia, and *G. sulphuraria* from North America. If Cyanidium strains are confirmed in NZ this would indicate a potential third migration event. The *G. sulphuraria* migration event is significant given the large distance between YNP and NZ, it is evidence for a long-distance dispersal mechanism for Cyanidiales or that there are unsampled “stepping stone” sites between NZ and the U.S.A. There are a number of volcanic and geothermal sites in the Pacific that are suitable environments for Cyanidiales to thrive. Interestingly, there have been no Cyanidiales strains documented in Hawaii [86,87], despite suitable environments.

Given the potential of *G. sulphuraria strains* in the development of Biotechnology (section 2.5), the identity of NZ type IV strains needs to be resolved. If the discrepancy between the *rbc*L and 18S rRNA sequence is repeatable, it would have greater implications for the evolution and phylogeny of Cyanidiales.

The current molecular data [26] is evidence that New Zealand is home to novel Cyanidiales strains, potentially with interesting phenotypes useful in development of biotechnology. However, to accurate identify novel phenotypes and resolve taxonomy of Cyanidiale requires a well round approach using both phenotypic characterisation and phylogenetic analyses.

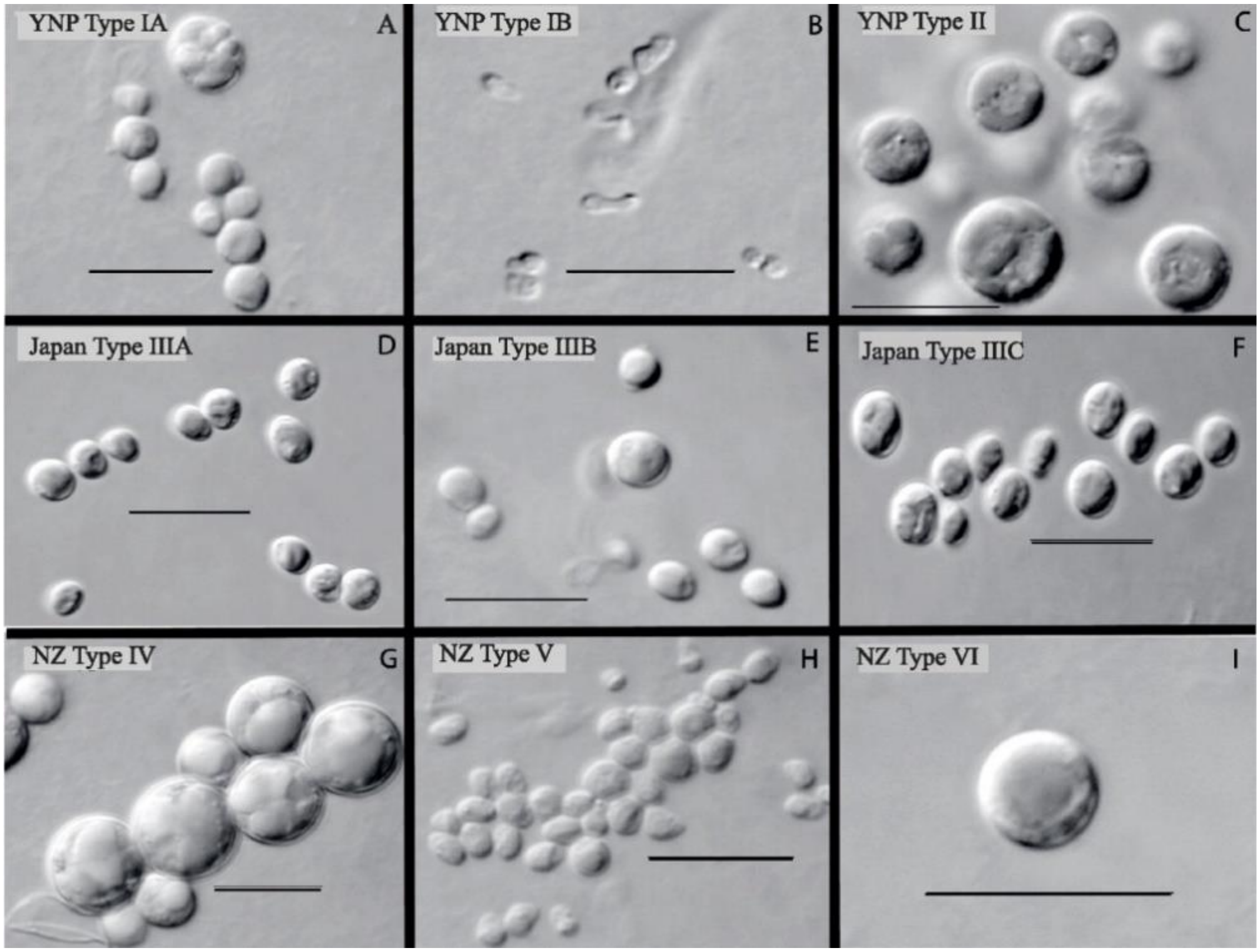


Figure 2.5: Microscope images of the different strains observed in Toplin *et al.* (2008). G,H and I are the New Zealand strains. B-YNP Type IB are considered to have true *C. merolae* morphotype. Sourced Toplin *et al.* (2008)

Table 2.2 Nearest BLAST relative for each different strain in Toplin *et al.* (2008) from YNP, Japan, and New Zealand for both *rbc*L and 18S rDNA. Sourced Toplin *et al.* 2008

Isolate source	Type	Nearest BLAST relative[a] *rbcL*	Nearest BLAST relative[a] 18S rDNA	% Variation within group[b]
YNP	IA/IB	*C. merolae* 10D (99)	*C. merolae* 10D (99)	0.38/1.9
	II	*G. sulphuraria* UTEX 2393 (99)	*G. sulphuraria* SAG 107.79 (98)	0.38/4
Japan	IIIA	*G. maxima* IPPAS507 (99)	*G. maxima* IPPAS P507 (100)	1/0
	IIIB	*G. maxima* IPPAS507 (96)	*G. maxima* IPPAS P507 (100)	1/0
	IIIC	*G. maxima* IPPAS507 (93)	*G. maxima* IPPAS P507 (99)	0
New Zealand	IV	*G. sulphuraria* SAG 108.79 (96)	*G. maxima* IPPAS P507 (99)	0.8/1
	V	*G. maxima* IPPAS P507 (93)	*G. maxima* IPPAS P507 (99)	1/1
	VI	*G. maxima* IPPAS P507 (91)	*G. maxima* IPPAS P507 (99)	0

[a] The numbers in the parentheses indicate the percent identity to the nearest GenBank relative.
[b] Results are shown as percent variation for *rbcL*/percent variation for 18S rDNA.

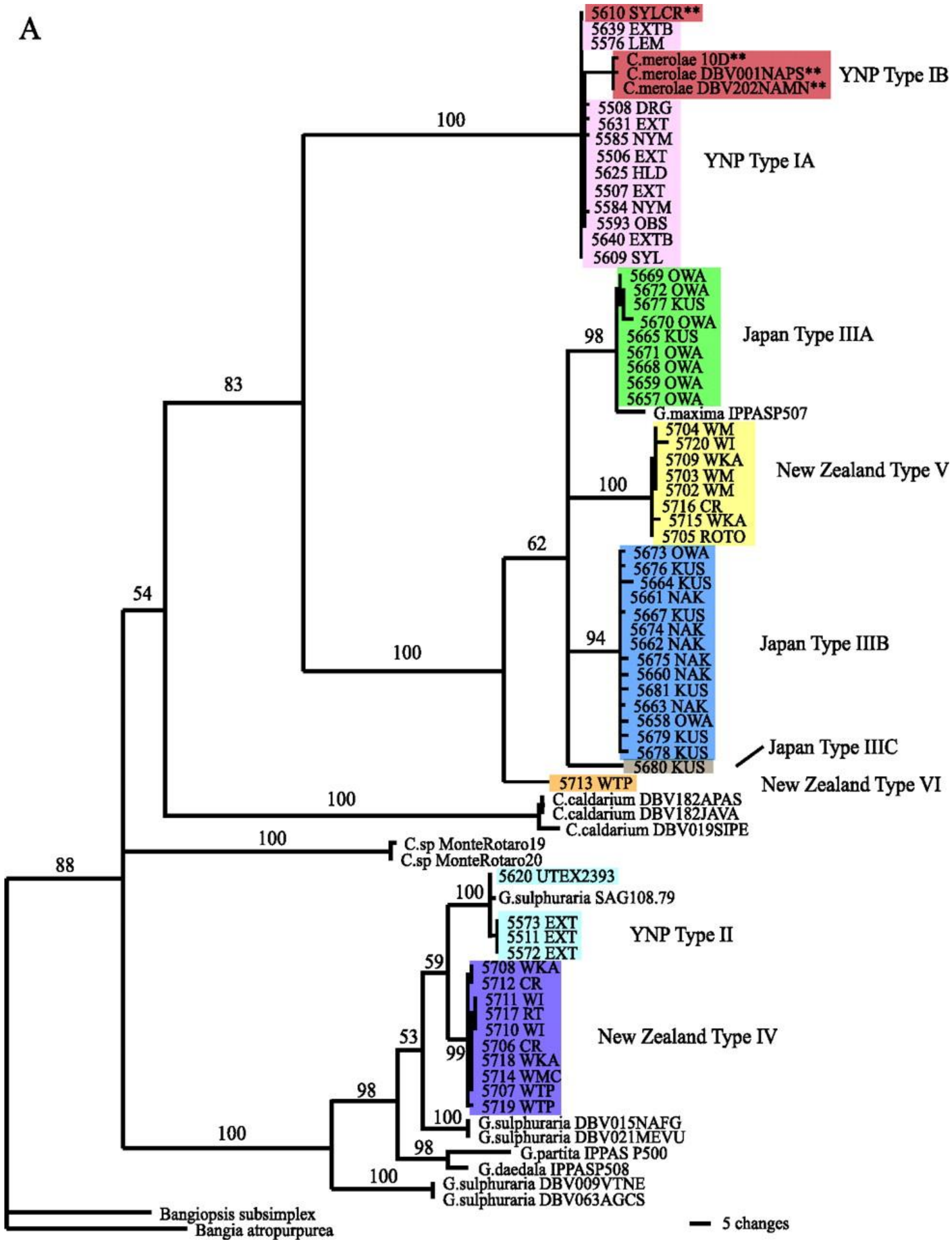


Figure 2.6: Phylogeny of strains from Toplin et al (2008). Inferred by maximum parsimony(MP) analysis of the *rbc*L gene sequences. Bootstrap values are shown on the branches. Sequences from the study are coloured with respect to their Type. Sourced from Toplin et al (2008)

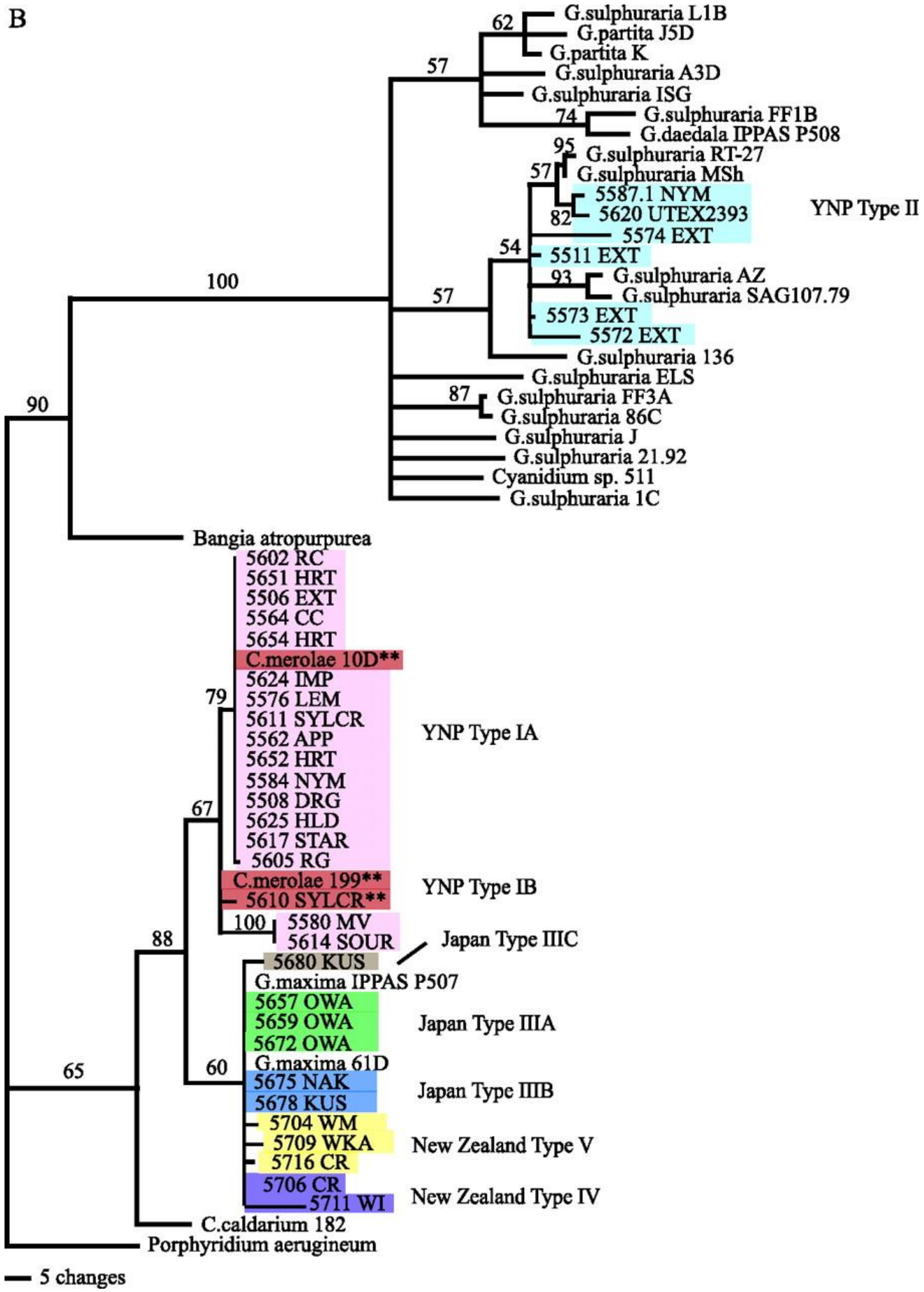


Figure 2.7: Phylogeny of strains from Toplin et al (2008). Inferred by maximum parsimony(MP) analysis of the 18S rRNA gene sequences. Bootstrap values are shown on the branches. Sequences from the study are coloured with respect to their Type. Sourced from Toplin et al (2008)

2.4 Cyanidiales in the Genomic Age

Whole genome sequencing is a powerful tool as it allows for the prediction of all possible proteins, detection of HGT and higher resolution of phylogeny.

At the start of this project, only the genome sequences of *C. merolae* 10D[42,43], *G. sulphuraria* 074W [44] and a draft genome of *G.phlegrea* DBV009 [45] were available/ published. Since commencing, the genome of *G.phlegrea* has been completed and the draft genomes of an additional 10 Cyanidiales strains have been published [46]. These genome sequences have given insight into the evolution of Cyanidailes, Rhodophyta and Archaeplastida. These genomes provide strong evidence that the genome reduction is significant in the evolution and diversification of Cyanidiales and Rhodophyta in general [76]. Additionally, it has provided strong evidence of HGTs from prokaryotes to Eukaryotes that incur fitness benefit [44–46]. The genome sequencing of *C. merolae* 10D also represented scientific milestones with the sequencing of its mitochondrial genome [88], plastid genome [89] and finally the nuclear genome [42], it was the complete algal genome sequenced. *C. merolae* 10D is also the first eukaryote to have genome sequenced with 100% coverage, from telomere to telomere on all chromosomes [43]. The genomes indicated that *C. merolae* was simplest known photosynthetic eukaryote and was comparable to the simplest non-photosynthetic non-symbiotic eukaryotes known, like *Saccharomyces pombe* and *Ashbya gossypii*[43]. The nuclear genome of *C. merolae* 10D is 16.5Mbp (megabase pairs) long across 20 chromosomes, which is longer than the approximately 12.5Mbp genomes of *Saccharomyces* spp. [42] ,however only 4775 protein-coding sequences were identified [43]. This is comparable to the simplest known non-symbiotic eukaryote *A.gossypii* which has a genome approximately 9Mbp long with 4726 genes [43,90]. The genome sequence also revealed that only 0.5% of genes contained introns and only 0.7% of transposable elements in the genome. Additionally, *C. merolae* 10D only had three copies of the rRNA gene set, a single cluster of Histone genes and a limited Cytoskeleton-related gene set, with no myosin gene found and an unexpressed actin gene. While this could be an indication of *C. merolae* representing a primitive morphology, it is more likely to be evidence of dramatic genome reduction and gene loss [42,43,76]. The mosaic of enzymes involved in the Calvin cycle were present and highly conserved when compared to *Arabidopsis thaliana*, [42] supporting that primary endosymbiosis was a single event before the divergence of Rhodophyta and green plants.

C. merolae 10D is an obligate photoautotroph and lacks a cell wall, while *G. sulphuraria* 074W was a facultative heterotroph which was able to utilize a wide range of carbon sources [91] and had a cell wall. Comparing the genomes of *C. merolae* 10D and *G. sulphuraria* 074W gave clear molecular reasons for the difference in observed morphologies. Initial comparisons used a large EST (Expressed Sequences Tag) data set from *G. sulphuraria* 074G [92] and the 8Mbp draft genome of *G. sulphuraria* 074W [57], until its completion of the genome [44]. The nuclear genome of *G. sulphuraria* 074w was shorter and more compact when compared to *C. merolae* and other Eukaryotes. The genome was 13.7Mbp long, with 6623 protein-coding genes and the coding sequencing making up 77.5% of the whole genome [44]. Unlike *C. merolae*, more than 50% of the coding genes in *G. sulphuraria* contained introns [44,57], and only 42% of the protein-coding genes found in *G. sulphuraria* had orthologs in *C. merolae*. This difference in protein sequences is the molecular basis for the differences in observed morphology. The genes involved in the biosynthesis of floridoside, an important osmolyte [93,94] and cell wall precursor in Red algae [95], were present in the *G. sulphuraria* genome but absent in *C. merolae* genome. The presence of floridosides had been biochemically observed in *Cyanidium* and *Galdieria*, but only found in trace quantities in *Cyanidioschyzon* [96–98]. The molecular basis for *G. sulphuraria* metabolic flexibility compared to *C. merolae* is more complicated. The *G. sulphuraria* genome coded several sugar kinases and carbon metabolic pathways, however a number of these kinases and pathways were found partial or completely intact in *C. merolae* genome, including complete metabolic pathways for glucose, mannose and galactose [57], however, there are a number of sugar transporters found in the *Galdieria* genome that are absent from *C. merolae* genome [44,57]. Membrane transport proteins seem to be key to *G. sulphuraria* metabolic flexibility and its high tolerance to stress factors like heavy metals and salinity, with about 5.2% of the coding genes coding for transporters[44]. Phylogenetic analysis indicates that many of these transporters have prokaryotic origins, sharing close homology with genes from extremophilic bacteria and thermophilic archaea [44]. This was some of the first strong evidence for HGTs from prokaryotes to eukaryotes that provided a fitness benefit but has been supported by additional genomic studies of Cyanidiales. The recent surveying of 10 novel draft Cyanidiales genomes and the three already discussed ones, indicate that on average about 1% of coding genes in Cyanidiales have been acquired via HGT[46]. All of this genomic data indicates that the evolution and diversification of Cyanidiales is the result of a combination of dramatic genome reduction and HGT driven by the selective pressure from the polyextreme environments of geothermal sites. A current theory is that the Cyanidiales common ancestor had a cell wall, and while it may have been heterotrophic, it later became an obligate autotroph after its migration into geothermal sites. The ancestral membrane transporters would not have functioned against the steep proton gradient of acidic environments and denaturing would probably occur at the

elevated temperatures. The sugar transports were not required for survival due to the acquisition of a photosynthetic plastid, so dramatic genome reduction occurred, leading to the loss of many non-functional genes and pathways in the Cyanidiales common ancestor. The "true" *Galdieria* lineage diverged approximately 1bya by acquiring functional transporters via HGT events, which allowed *Galdieria* to establish an ecological niche in drier, lower light environments, e.g. endolith environments. Given the close relation between *Cyanidioshyzon* and *G. maxima*, which each form a monophyletic clade with *Cyanidium* (fig. 2.3), it suggests that the absence of a cell wall and the floridoside pathway in *Cyanidioschyzon* strains is not an ancestral feature. Rather it is the more recent consequence of mutations and genome reduction while occupying osmotically stable environments, as true *C. merolae* ecotype has only been observed in hot springs and geothermal water (ref). In this theory, *G. maxima* acquired the appropriate genes for heterophic growth via HGT independently of the main *Galdieria* lineage. It is difficult to estimate the timeline for HGT events in the *G. maxima* lineage, as there is evidence that certain *C. caldarium* strains are not strict obligate photoautotrophs and are capable of heterotrophic growth under certain conditions [59]. Currently, there are no genome sequences available for either a *G. maxima* strain or a *C. caldarium* strain, which could be critical evidence to support this theory. In this regard, the potential presence and subsequent cultivation of *C. caldarium* and *G. maxima* strains in New Zealand presents an opportunity to contribute to this narrative as part of this thesis. If the above theory is correct, it would be expected that the genome of *G. maxima* would share closer homology to *C. merolae* than to the *Galdieria* genomes. Additionally, if the strain is a facultative heterotroph it would be expected that the related genes, such as sugar transporters, would have been acquired from HGT, but would not share homology to the orthologs in the other *Galdieria* genomes as it was acquired separately.
The molecular evidence that genome reduction and HGTs are significant in the evolution and diversification of Cyanidiales is strong and mounting [44–46,57,76,99,100]. However, caution is advised when extrapolating this data to other Eukaryotic lineages, especially multicellular and/or sexual lineages. Genome reduction in early Rhodophyta is supported by the genome sequences of other Rhodophyta, such as *Chondrus crispus* [100] and *Porphyridium purpureum* [101]. So evidence suggests that genome reduction was important in the evolution of Rhodophyta, but is less significant in the evolution of Green algae and Land plants [76].
At the time of writing, the presence of HGTs in Eukaryotes, outside of the context of endosymbiosis and pathogenic events, is highly controversial [102–104]. The morphology and environments of Cyanidiales make the order potentially unique or rare among Eukaryotes for HGTs. Cyanidiales are unicellular and asexual, so HGT events can occur more frequently and

are guaranteed to be past vertically. Additionally, the geothermal sites they occupy will be causing significant selection pressures and is home to other specialized organisms like thermophilic archaea, which provides an expanded gene pool. Therefore, this data may not apply to multicellular, sexual eukaryotes where the HGT, like as all mutations, needs to be incorporated into the gametes to be inherited, reducing frequency. Similarly, the frequency of HGTs in mesophillic unicellular, asexual Eukaryotes will be significantly less as selection pressures are comparatively lower. For example, the unicellular *P.purpureum*, in the genes from HGT are identified with less certainty and those with certainty seem to have occurred in an earlier ancestor [101]. As genome sequencing technology develops, costs will reduce, and more genomes will be sequenced. In time we may resolve the evolution of Cyanidiales and HGT in Eukaryotes.

2.5 Development of Biotechnology and Cyanidiales

2.5.1 Cyanidiales potential in Biotechnology

As polyextremophile alga, the Cyanidiales are promising candidates for the development of microalgae biotechnology. Extremophiles are important in large-scale commercial applications of microorganisms, as the imposed culturing conditions are generally non-conducive to the growth of most organisms and therefore reduce the incidences of contamination [14,21,47]. Many commercially used microalgae species are extremophiles, examples include; the alkaliphilic Spirulina (*Arthrospira maxima* and *A.platensis*) which is used as a food product and to produce pigments like phycocyanin[47,105,106], and the halophile *Dunaliella salina* which is used to produce B-carotene[14,47]. As a facultative heterotroph with tolerance to heavy metals and salt, *G. sulphuraria* is the most promising biotechnology candidate of the Cyanidiales and biotechnology development is one of primary driving factors for research into *Galdieria* spp. [47]. *Galdieria* has been studied for use in a wide range of applications including phycocyanin production, heavy metal recovery and treatment of effluent [21,47,107].

2.5.2 Phycocyanin production with *Galdieria sulphuraria*

Phycocyanin is a blue phycobiliprotein, a photosynthetic pigment found in cyanobacteria and some red algae, like the Cyanidiales[48,108]. Phycocyanin is a high value microalgae-derived product, commercially sourced from cultivated Spirulina[105,106] and is used as a stable natural blue food dye, an natural antioxidant and a clear fluorescent marker for biomedical assays[106,107]. However, there is plenty of research in using *Galdieria sulphuraria* 074G for commercial phycocyanin production. While the yield of phycocyanin per gram of dry weight is around ten times less in 074G compared to Spirulina [107,109] the number of advantages in using *G. sulphuraria* 074G over Spirulina.

1) *G. sulphuraria* 074G can be grown heterophically and maintains production of phycocyanin[91,110]
2) *G. sulphuraria* 074G can grow heterotrophically on a wide range of substrates[91]
3) The phycocyanin from Cyanidiales is more thermally stable[111,112].

As such it is possible to grow heterotrophic *G. sulphuraria* culture where the increased growth rate and cell density means Phycocyanin production rate can be up to 100 times higher than Spirulina[105,108,113,114]. The cost of supplying a carbon source can be offset by growing the culture using a waste product such as bakery waste[115] and the thermostable phycocyanin[111,112] is more valuable as it can be used as a dye in a greater range of applications. There is a strong possibility that this Cyanidiales biotechnology will become economically viable and replace Spirulina[107].

2.5.3 Metal recovery with *Galdieria sulphuraria*.

The highly selective and rapid metal recovery capabilities of *G. sulphuraria* have been documented since the late 1980's [116], where it was shown that *G. sulphuraria* selectively recovered 99% of the copper from acidic bacterial leachate resulting in 7% Cu DW in 96hrs. Although only recently has the topic regained interest [47] with new research in the last 5 years. It has been shown that *G. sulphuraria* can recover 90% of rare earth metals like; neodymium(Nd), dysprosium(Dy) and lanthanum(La) from acidic aqueous solutions[117]. *G. sulphuraria* is capable of selectively recovering 100% of precious metals like gold(Au), palladium(Pd) and platinum (Pt) from aqua regia-based metal wastewater in an hour [118]. This metal recovery has a number of applications. It could be used for remediation of polluted sites like acid-mine waters [116,119] or as suggested in [120] removal of radioactive cesium from Fukushima. It could be used for recycling, for example, electronic waste contains a lot of rare earth metals like Nd that are difficult to recover, or recovery of catalysts like platinum from industrial wastewaters.

2.5.4 Single-step wastewater treatment with *Galdieria sulphuraria*

The research and development of treating effluent or urban wastewater utilizing *G. sulphuraria* began recently [47,121], but is a promising topic with research going from laboratory scale[121,122] to field scale, with 700L batch photobioreactors [123] within a few years. The core concept is that in a photobioreactor, a *G. sulphuraria* culture grown in the primary effluent would remove all the nutrients and organic compounds (measure as BOD, biochemical oxygen demand) through mixotrophic growth, and the acidic conditions would sterilise pathogenic bacteria [122]. If practical, this would reduce secondary treatment of effluent from a multistep process with multiple microorganisms (fig. 2.8), to a single-step, single organism process, and the biomass could be converted into a valuable product such as biofuel[124,125]. The 700L fed-batch closed photobioreactor [21] used in the field scale trials; did not require additional heating or nutrient/organic supplementation, reduced the BOD, nitrogen and phosphate content in the primary effluent to the discharge standard within 3 days and the total coliform count was reduced below detection limits in a day [21]. The authors did note that a practical limitation could be the pH adjustment, the primary effluent needed to be acidified for optimal *Galdieria* growth and the treated effluent had to be neutralised before discharge. Another practical concern is temperature control, while the field reactors maintained temperature between 30-50°C with solar heat alone [21,123], these reactors were set up in a wastewater treatment plant in New Mexico. For treatment plants in colder or more temperate climates, active heating would be required but could be done using the biogases generated in the process. The Christchurch municipal wastewater treatment plant already captures methane and uses it in power generation to keep digesters at optimal temperature (fig. 2.8) so could be practical in New Zealand.

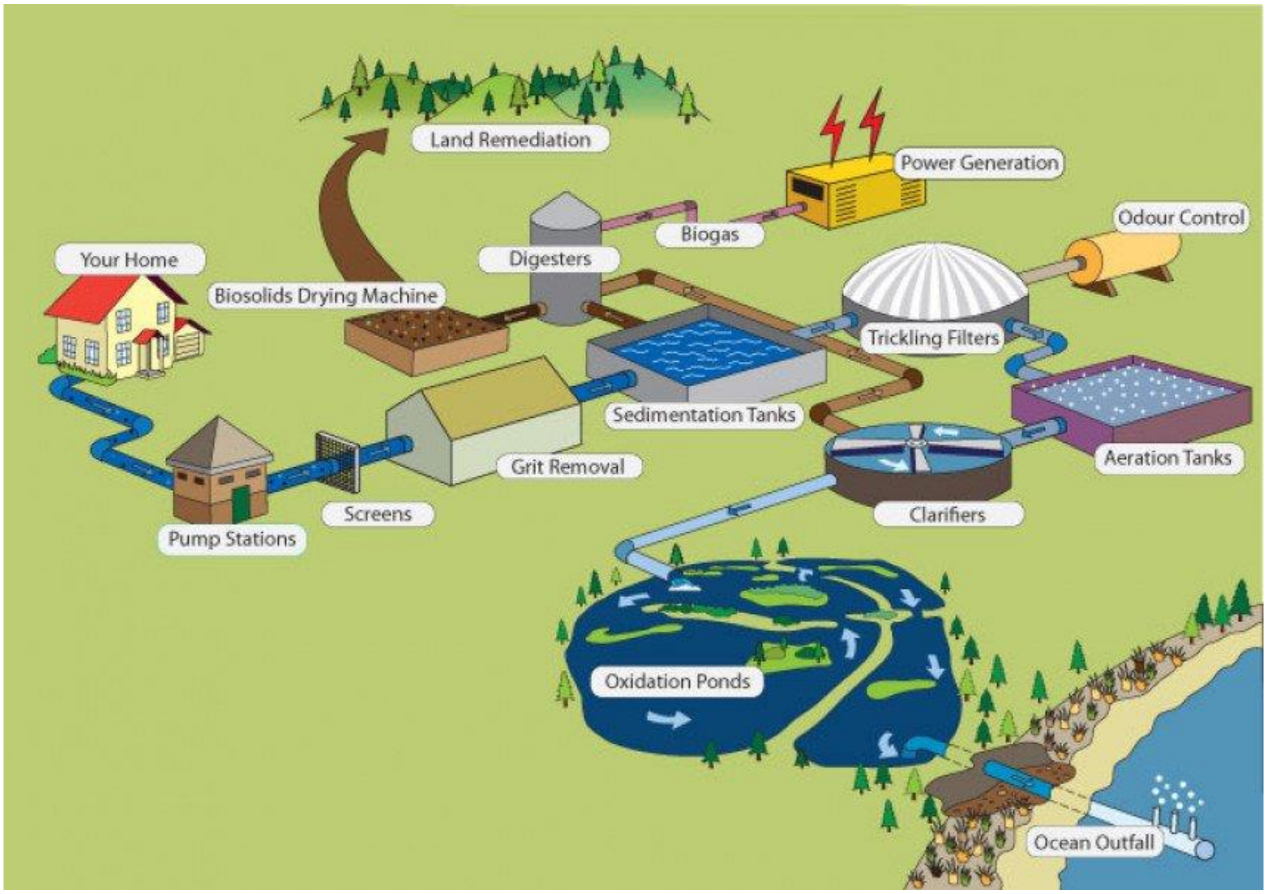


Figure 2.8: Systematic Diagram of the Christchurch municipal wastewater treatment plant (Sourced from ccc.govt.nz accessed 03/2020)

2.5.5 Sequestering flue gases with Cyanidiales

In terms of the MBIE project and my Master's thesis, there are some challenges sequestering CO_2 from flue gas, but Cyanidiales in theory can overcome these challenges. Flue gas composition, besides the organic gases, includes high levels of sulphur oxides (SO_x) and nitrogen oxides (NO_x). These gases are innately toxic to many organisms and produce acidic compounds when reacted with water, which lowers the pH of the medium [126–128]. The Cyanidiales, as acidophiles are tolerant to sulphuric compounds and high CO_2 concentration, therefore are ideal candidates for the development of this biotechnology. However the research on Cyanidiales in this context is limited [127,129–132]. It has been documented that the growth of *C. caldarium* was unaffected by 40% stimulated flue gas [129] and that *C. caldarium* exhibited growth in flue gas concentrations inhibitory to *Chlorella vulgaris* [130]. In a comparison between *C. merolae*, *C. caldarium* and *G.partita*, *Galdieria partita* had the highest tolerance to SO_2 [131].

In summary, the development of biotechnology using Cyanidiales is still in its infancy and is yet to reach large-scale commercial use. However, progress from laboratory work to field trails is promising. This field of research is heavily focused on *G. sulphuraria* strains, due to the species' metabolic flexibility and stress tolerance, compared to other Cyanidiales. This emphasises the importance of determining if NZ type IV strains (section 2.3 table 2.2) are representatives of *G. sulphuraria*, as then the above biotechnologies can be more easily developed in New Zealand. Failing which, it would be ideal to isolate other *G. sulphuraria* representatives from New Zealand.

2.6 Objectives and hypotheses

The objective of this MSc thesis was to sample and isolate Cyanidiales from New Zealand geothermal sites in the Taupō Volcanic Zone (TVZ) and gather genetic, morphological and physiological data, in order to describe the strains and resolve taxonomy.
Evidence based on the phylogeny of *rbc*L sequences [26,32] indicates that native New Zealand Cyanidiales strains descended from at least two lineages; the *Galdieria sulphuraria* lineage, and the *Galdieria maxima* lineage, represented by NZ type IV strains and NZ type V & VI strains respectively [26]. Meanwhile, the phylogeny of 18S rRNA sequence [26] suggest that NZ Cyanidiales belong to a single lineage, the *G. maxima* lineage. This conflict is due to mismatch between the *rbc*L phylogeny and the 18S rRNA phylogeny of NZ type IV strains[26], and there are a number of potential reasons for this conflict including experimental error (section 2.3). Therefore it is important to use correct isolation techniques and determine taxonomy using a combination of phenotypic and phylogenetic methods as more data a identification of experimental error.
There is also evidence that NZ Cyanidiales may include strains from the *Cyanidium* lineage as well, based on molecular data from the plastid 16S rRNA sequence from a NZ Cyanidiales isolate (strain NZ4) [27]. However, 16S rRNA molecular data has historically been unreliable in the identification of Cyanidiales in YNP, where the plastid 16S rRNA genes of *Galdieria* strains showed high pairwise identity (>90%) to *Cyanidium* hence were misidentified [26,83,84]. This could have occured with NZ4 strain and it is a misidentified *Galdieria* strain potentially due to the gene being highly conserved within Cyanidiales.

In order to answer these questions, I will need to sample and isolate several Cyanidiales from different geothermal sites in New Zealand. I will then characterise and taxonomically classify these isolates using a combination of physiological, morphological and phylogenetic evidence. Traditionally Cyanidiales genera have been differentiated by key phenotypic characteristics (table 2.1) [30,31] which in conjunction with molecular data can be used to resolve and confirm taxonomy further. For Cyanidiales, commonly *rbc*L and/or 18S rRNA are used to determine phylogeny data however three plastid protein gene dataset [133] and a mixed 5-gene dataset [29] have been used to more accurately assess the diversity of Cyanidiales lineages. Such characterization has not been done on these New Zealand strains, hence there is uncertainty in their taxonomy.

In summary, the hypotheses of this thesis are:

A) New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* lineage

B) The two New Zealand "*Galdieria*" lineages (*G. sulphuraria and G. maxima*) can be differentiated by phenotypic characteristics and phylogenetic analysis.

C)The plastid 16S rRNA gene of my isolates will have a close similarity(>99% pairwise identity) to *C. caldarium*.

To address these hypotheses, I will:

1. Enrich and isolate Cyanidiales from a wide range of NZ geothermal ecosystems;
2. Confirm isolation of at least one of each NZ type IV and V strains using rbcL and 18S rRNA sequences; and
3. Characterise and compare the phylogenetic, morphological, ecological and physiological traits of Cyanidales isolates.

In addition, I will:

4. Use the physiological characteristics of isolates to assess suitability for biotechnology application

This thesis will not only contribute to the MBIE Smart Ideas project by increasing the number and diversity of Cyanidiales strains available to the CAPE/Scion researchers, but will also contribute data which is useful for the larger scientific community. The data will provide evidence that can determine which Cyanidiales lineages are present in New Zealand, aiding in answering questions and hypotheses about the diversity, dispersal and evolution of Cyanidiales both in New Zealand and globally. For example, the evidence suggests that there were at least two separate Cyanidiales colonization events to NZ, *G. maxima* through Eurasia and *G.*

sulphuraria from the Americas [26,32] which is either indication of a long-distance dispersal mechanism or unsampled sites between NZ and the Americas. Although the American *G. sulphuraria* lineage needs to be confirmed before conclusion can be drawn. The isolation of Cyanidiales strains from New Zealand will likely lead to the discovery of novel strains and species, potentially with novel characteristics. Which would give opportunities to study evolutionary pathways that developed those novel ecotypes. Additionally this thesis will contribute New Zealand native Cyanidiales strains that could contribute to the development and research of other biotechnologies in New Zealand such as, the production of pigments like phycocyanin [107,114], feedstocks [12] and bioremediation [14,120,121].

Chapter 3. Sampling, Isolation and Phenotypic Characterization of NZ Cyanidiales

3.1 Introduction

3.1.1 Background

Cyanidiales includes two families Cyanidiaceae and Galdieriaceae[30]. Cyanidiaceae includes *Cyanidium* spp., *Cyanidioschyzon* spp. and the misnamed *Galdieria maxima*[38]. Galdieriaceae includes *Galdieria sulphuraria* and *G. phlegrea*[38,75]. Key diagnostic features of the genera are listed in table 2.1.

Cyanidiales species live in various environments, however most strains are obligate acidophilic thermophiles, requiring acidic (<pH 4) environments with temperatures above 40°C [30,31]. As such, many strains are restricted to, and thrive in, sulphuric geothermal areas. There are a number of such sites in New Zealand within the Taupō Volcanic Zone (TVZ). The TVZ is in the North Island, running North-East from Mount Ruapehu to White Island/Whakaari, through Taupo and Rotorua (fig. 3.).

Cyanidiales have been observed at many sites in the TVZ such as Waimangu Cauldron [24], White Island/Whakaari [27] and Waitapu [26,134]. However, the isolation, characterisation and identification of Cyanidiales in NZ and the TVZ has been limited due to a lack of research. To date, in New Zealand, a *Cyanidium caldarium* (NZ4) has been identified by plastid 16S rRNA gene similarity [27], a *Galdieria sulphuraria* (NZ type IV), and two *Galdieria maxima* strains (NZ type V and VI) have been identified with *rbc*L and 18S rRNA gene phylogeny [26]. The identity of the *G. sulphuraria* strains is uncertain due to conflict between the *rbc*L and 18S rRNA gene phylogeny[26]. However, characterisation has been limited to observed cell morphology and measuring the conditions of the sampling environment. Cell morphology can be used to determine some genera, like *Cyanidioschyzon*, but lacks the resolution to accurately distinguish between *Cyanidium* and *Galdieria* spp. due to similar cell morphology (table 2.1).

Phenotypic characterisation to accurately identify the genus and the species of a Cyanidiales sample requires characterisation of a number of morphological, ecophysiological and biochemical traits/features. This includes pH growth range, temperature growth range, nitrogen source utilization, facultative heterotrophy and presence of trienoic fatty acids. This in-depth characterisation can determine if strains have a novel phenotype, which would not be revealed by phylogenetic data, and it can help determine a strains suitability to certain biotechnology applications. For example, the bioreactor may require certain metabolism and/or will need to be maintained at a specific temperature and pH range.

The different genera of Cyanidiales are adapted to different micro-environments within geothermal areas. For example, *Cyanidioschyzon* spp. are obligate photoautotrophs and do not tolerate osmotic stress, therefore *Cyanidioschyzon* spp. are restricted to hot springs with light. Whereas *Galdieria* spp. are facultative heterotrophs with high tolerance to desiccation [31,36] , hence can live in drier and darker environments such as acidic geothermal heated soil. Therefore to maximise the chances that diverse range of Cyanidiales strains are isolated, sampling sites need to be geographically and environmentally distinct. Also a wide range of enrichment conditions need to be used.

Figure 3.1: Taupo Volcanic Zone a region of geothermal activity, North Island, New Zealand. Sourced from (http://www.explorevolcanoes.com accessed 05/2019)

3.1.2 Aim of this section

The Aims of this chapter are to: 1) sample and isolate Cyanidiales from New Zealand geothermal sites in the Taupō Volcanic Zone and 2) generate morphological, physiological and biochemical data in order to describe the strains.

This will determine the taxonomy of the isolates, identify novel phenotypes and increase the number New Zealand isolates available for biotechnological developments.

This would also address two of the previously stated hypotheses, which are:

A) New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* lineage
B) The two New Zealand “*Galdieria*” lineages (*G. sulphuraria* and *G. maxima*) can be differentiated via phenotypic characteristics and phylogenetic analysis.

To achieve this, I will take several Cyanidiales samples from micro-environments across multiple sites within the TVZ. As there is evidence that the NZ population is diverse, with several different genera and species, it is important to take multiple samples that are both environmental and geographically distant. This increases the probability of isolating not only examples of previously described strains, but also novel undescribed strains. The sampling sites were; Craters of the Moon, Taupo; Kuirau Park, Rotorua; and Parariki Stream, Rotokawa. These sites were chosen as Cyanidiales had been previously observed at these sites [25,26], the sites were geographically distant and have different geothermal features. Additionally, we had permission from the respective Iwi and landowners to sample and use isolates for scientific and commercial applications. I will generate axenic cultures using the streak plate method, an antibiotic treatment that I developed and a medium that I modified for this purpose. Four of the axenic cultures were chosen for further characterisation, which included experiments to; determine temperature growth range, pH growth range, nitrogen source utilization, heterotrophy and lipid profile.

3.2 Methods

3.2.1 Sampling

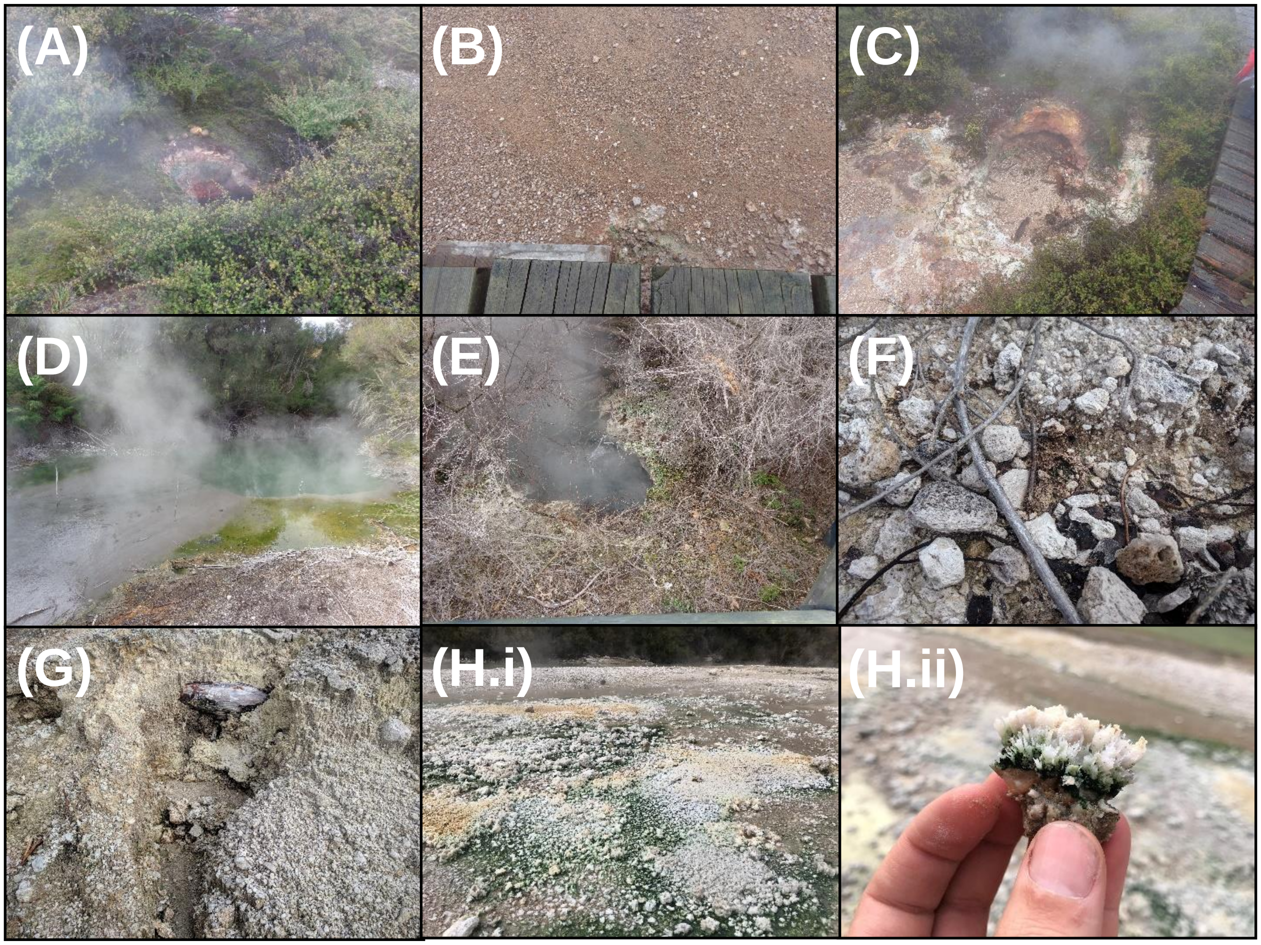


Figure 3.2*:* The environmental sampling sites with the TVZ **A-C)** Craters of the Moon **A)** Sample CM1.1 **B)** Sample CM1.2 **C)** CM1.3. **D, E)** Kuirau park KP1.1 and KP1.2 respectively. **F-H)** Rotokawa **F)** RK1.1 **G)** RK1.2 **H.i)** RK1.3 **H.ii)** The silica rock formations with green cyanidiales biomass.

The sampling trip to the Taupō Volcanic Zone occurred from the 22/04/2019 to the 26/04/2019. Sampling sites were: Craters of the Moon, Taupo; Kuirau Park, Rotorua; and Parariki Stream, Rotokawa. Cyandiales samples were identified in the field by green biomass at geothermal sites known to be acidic. Samples were aseptically collected into sterile 50mL Centrifuge tubes, using a sterilized spatula or the tube directly. Samples were kept in the dark at room temperature for 10-14 days before enrichment. In total, eight samples were taken, which includes three samples from Craters of the Moon, two samples from Kuirau Park and three samples from Rotokawa (Table 3.1 and Figs. 3.2). On-site, a digital thermocouple was used to measure temperatures and a GPS was used to determine the NZTM2000 coordinates (Table 3.1). To determine environmental pH, an additional sample of the substrate was taken. For dry substrates (soil/sinter) 2mL of the substrate was placed in a sterile 15mL centrifuge tube and MillQ water was added to the 10mL mark. The sample was agitated by hand and the substrate was allowed to settle. The pH of the solution was measured using a pH meter (Table 3.1). The pH of the water samples was measured directly with the pH meter.

3.2.2 Enrichments

Media for enrichments were prepared in 60mL serum bottles, with 30mL of V4 [23,135] media (section 6.1) at pH 2.5, 1.5 and 0.7. Bottles were sealed and the headspace was enriched with 40mL of CO_2 using a syringe, then autoclaved. For environmental samples collected from streams or waterbodies, water was used as a direct inoculum after brief agitation to resuspend settled cells. For each environmental soil or sinter sample, a small amount of substrate, approximately 1-2mL, was placed in a 15mL centrifuge tube and filled to 10mL with sterile water and agitated by hand. Each sample was allowed to settle briefly and the supernatant was decanted into a fresh Centrifuge tube For each sample, 0.5mL of the respective supernatant was aseptically injected into nine sealed serum bottles, three at each pH value. The serum bottles were then incubated at 45°C, 55°C or 60°C under constant illumination using fluorescent lights. Therefore, each sample was enriched under nine different conditions, across three temperatures and three pH values. Enrichments were incubated for 28 days. Growth of Cyanidiales was verified by observing a green colour change in the medium and viewing cells under a light microscope (section 3.2.4). Samples that had no successful enrichments were re-attempt using V4 medium at pH 4.0 and 5.0, or Modified Allen's medium [136] (recipe section 6.1) at pH 2.5 incubated at 45°C.

3.2.3 Isolation

Axenic cultures were generated using the streak plate technique. Plates used in the final isolation were used MAll-Ht plates, 1.5% Phytagel Modified Allen's medium [136] supplemented with 3gL^{-1} glucose and 3gL^{-1} glycerol (see section 6.1 for recipes). Initial streaking from enrichments was treated with antibiotics to remove bacterial contamination. In an Eppendorf tube, a 0.5mL aliquot, from the Enrichment serum bottles, was made up to contain either 25mgL^{-1} of Ampicillin or a mixture of 25mgL^{-1} Ampicillin, 12.5mgL^{-1} Vancomycin and 12.5mgL^{-1} Erythromycin, and incubated for an hour at room temperature. The tubes were then vortexed

and 100µL was pipetted onto the plate for streaking. Cultures were considered axenic cultures after three successive passages, once the colony morphology was stable. This was verified by observation under a light microscope (section 3.2.4).
Generating axenic cultures was attempted with other isolation techniques and different solid media. Serial dilution by extinction was trailed using 10ml of V4 media (section 6.1) in 25mL anaerobic cultures tubes. Other solid media attempted using the streak plate technique included 2xV4 medium and Modified Allen's from Minoda *et al.* 2004. To remove bacterial contamination exposed to UV light was tested and antibiotics, including those listed above and tetracycline were trailed individually and in combination to determine which

3.2.4 Microscopy

Microscopic observations were undertaken using an Olympus CX43 light microscope. Cells were observed under the 40x and 100x Oil immersion objectives with phase contrast and bright field.
Cells were immobilised by placing 10µL cell suspension on a thin block 0.2% (w/v) agar on the slides.

3.2.5 Nitrogen Utilization

Media for determining nitrogen utilization were prepared in 60mL serum bottles, with 30mL of 2xV4 media at pH 2.5 (H_2SO_4). Media was prepared with either 0.8gL^{-1} NH_4Cl, 1.6gL^{-1} KNO_3 or 0.8gL^{-1} urea as the sole nitrogen source and a nitrogen-free 2xV4 medium was prepared for negative controls. The media for the two yellow isolates, RK1.3 and CM1.3Y, were supplemented with 5mM of glucose to ensure growth. Bottles were sealed and the headspace was enriched with 60mL of CO_2 using a syringe, then autoclaved.
Biomass of the four selected isolates were harvested from MAll-HT solid media and suspended in nitrogen-free 2xV4 media in sterile 50mL Centrifuge tubes. The harvested cells were washed thrice this was done by repeated centrifugation (5 minutes, 5000x*g*), followed by removal of supernatant and resuspension in the sterile nitrogen-free medium. The serum bottles were inoculated with 0.2mL of the washed cells. The experiment was done in triplicates for all nitrogen sources, for the four isolates. Growth was qualitatively determined by visual inspection at endpoints of seven and 21 days at 45°C, under constant illumination.

3.2.6 pH Range

The medium for enrichments was prepared in 60mL serum bottles with 20mL of 2xV4 medium and pH was adjusted with H_2SO_4 or KOH, to obtain final pH values of; 0.0; 0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.5; 5.0; 6.0; 7.0. The media for the two yellow isolates, RK1.3 and CM1.3Y, were supplemented with 5mM of glucose to ensure growth. Bottles were sealed and the headspace was enriched with 40mL of CO_2 using a syringe, then autoclaved. Additional bottles were made at each pH value for non-inoculated controls. Prior to inoculation, pH of non-inoculated controls was measured using a pH meter and litmus paper. The bottles with the pH values of 7.0, 6.0 and 5.0 showed a slight drop in pH , decrease of 0.1-0.2 pH unit, and bottle 0.0 showed true pH value of 0.2. For all other bottles, pH values were within 0.1 pH unit of the initial medium pH. Biomass of the four selected isolates was harvested from MAll-HT plates and suspended in sterile water, containing 0.5gL^{-1} of $MgSO_4.7H_2O$ to reduce osmotic stress, in sterile 50mL centrifuge tubes. The serum bottles were inoculated with 0.2mL of the suspended cells. The experiments were undertaken in triplicate for all pH values for the four isolates. Growth was qualitatively determined by visual inspection at endpoints of 7 and 21 days at 45°C, under constant illumination. Post incubation pH values of both the controls and inoculated samples were measured using acidic (0.0-6.0) and general (0.0-14.0) litmus strips.

3.2.7 Temperature

The growth temperature range was tested using MAll-HT spread plates. Biomass of the four selected isolates was harvested from MAll-HT plates and suspended in 1mL of sterile V4 medium, in Eppendorf tubes. Plates were divided in half, with each side being a single isolate. 50µL of cell suspension was used for each half and carefully spread to avoid mixing. All isolates were grown at room temperature (23°C), 39°C, 45°C and 50°C, under constant illumination. The plates for the yellow isolates were also grown at 54°C and 60°C in the dark. The experiment was done in duplicates for all temperature values, for the four isolates. Growth was qualitatively determined by visual inspection after incubation for 14 days.

3.2.8 Heterotroph growth and carbon sources

Media for determining heterotrophic growth was prepared in 120mL serum bottles, with 60mL of 2xV4 media at pH 2.5 and supplemented with either 5gL^{-1} of glucose or 10mL of 50% Glycerol solution. Bottles were sealed and then autoclaved. Biomass of the four selected isolates was harvested from MAll-HT plates and suspended in 5mL of V4 medium, in 15mL Centrifuge tubes. The bottles were inoculated with 1.0mL of the cell suspension. Bottles were covered with aluminium foil to ensure cultures remained in the dark. Incubation was for 28 days at 45°C. The experiment was done in duplicates for both carbon sources, for the four isolates. Growth was qualitatively determined by visual inspection after incubation.

3.2.9 Fatty acids as methyl esters (FAME)

Fatty acid methyl esters (FAME) were prepared from biomass without lipid extraction and analysed as described by Svetashev et al[137]. The base reagent is made up of 1% metallic sodium in methanol. A loopful (2-5mg) of biomass is suspended in 200-300µL of the base reagent and then heated for 15mins at 100°C. After this, an equal volume of 5% HCl in methanol is added and the mixture is heated for 15 minutes at 100°C. Two 200µL aliquots of hexane are used to extract the FAMEs. The hexane phases are combined and the solvent is evaporated in a stream of inert gases, then 10µL of chloroform is used to resuspend the residue. This solution is used for analysis by GC-MS.
Fatty acid methyl esters identification for each isolate were performed on an Agilent 7890B gas chromatograph (Agilent, USA), equipped with FID and 30 m, 0.25 mm i.d. DB-WAX UI capillary column (Agilent, USA). Helium was used as the carrier gas; the split ratio was 1:30. Separation temperature was 195°C. Fatty acids were identified by the use of reference compounds, known ECL values, and by GC-MS. GC-MS analyses were performed on a GC-MS-QP2010 Ultra (Shimadzu, Japan).

3.2.10 Polar lipids and Thin layer chromatography

Total cellular lipids were extracted from the biomass using the method described in Bligh and Dyer 1959 [138]. One-part algal biomass, one-part chloroform and two parts methanol are homogenised for 2 min in a blender. Then another one-part chloroform is added to the mixture and homogenised for another 30 sec. The mixture is filtered through Grade 1 filtration paper under vacuum. The filtrate was allowed to separate and the methanol:water portion was removed.

Polar lipids were separated by one-dimensional thin-layer chromatography(1D-TLC) to allow comparison between the four isolates and, for better resolution, two-dimensional thin-layer chromatography (2D-TLC) using HPTLC silica gel 60 (10 x 10 cm, Merck). For the 1D-TLC the single direction is up, developed in chloroform:methanol:water (65:25:4, by volume). For the 2D-TLC the first direction (up) is developed in chloroform:methanol:water (65:25:4, by volume), and the second (left) in chloroform:methanol:acetic acid:water (80:12:15:4, by volume). Primary amines were detected with 0.2% ninhydrin in absolute ethanol. Phospholipids were detected with the molybdate spray reagent [139]. Glycolipids were detected with 0.5% anthrone in toluene. Another duplicate plate was charred by spraying with 10% v/v H_2SO_4 in methanol and incubating at 120°C for 30 minutes, in order to show total lipids.

3.3 Results

3.3.1 Sampling and enrichments

Eight Cyanidiales samples were sampled from three geothermal sites in the TVZ and enriched at different temperatures and acidities. This was to maximise the possibility of sampling and enriching different Cyanidiales strains. Of the eight samples collected, six samples had at least one successful enrichment, most were at 45°C and pH 2.5 (see table 3.1). CM1.3 the successful enrichment was at 45°C and pH 0.7. RK1.1 and KP1.1 did not have any successful enrichments even after attempts with Modified Allen's medium and higher pH values. No growth was observed in the 60°C enrichments nor were Cyanidiales cells observed under the microscope. The 55°C enrichment bottles had no observable growth after the 28 days but intact Cyanidiales cells could be observed under a microscope for samples CM1.3, RK1.2 and RK1.3 at pH 2.5. These three enrichments were concentrated via centrifugation and sub-sampled into new serum bottles with V4 medium at pH 2.5, then incubated at 45°C. Only the CM1.3 sub-samples showed observable growth. A total of seven of the enrichments were chosen for isolation, one from each sample that successfully enriched and the CM1.3 at 55°C enrichment (Table 3.1)

Table 3.1 Summary of sample sites, successful enrichments and final isolates

Location	Sample Number	GPS	Description	Temperature (°C)	pH	Successful enrichments	Isolates
Craters of the Moon	CM1.1	E1867007 N5717808	Small fumarole. Sample collected below lip	44.7	2.9	45°C (pH 2.5)	CM1.1G
	CM1.2	E1866990 N5717786	Heated soil, near the boardwalk. Green top layer	35-38	3.3	45°C (pH 2.5)	CM1.2G; CM1.2Y
	CM1.3	E1866896 N5717854	Large steam vent. Lots of steam. Sample at lip	50-60	2.7	45°C (pH 0.7)	CM1.3G; CM1.3Y
						55°C (pH 2.5)	CM1.355
Kuirau Park	KP1.1	E1884339 N5774702	Geothermal heat water sample collect at end of stream near auxiliary pool	51.8	6.5	N/A	N/A
	KP1.2	E1884184 N5774595	Sinter formation around a boiling spring (98°C) Cyanidiales on the top of the lip of sinter, most tree branches near spring branches were bare	50-55	2.3-2.5 (Soil) 7.5(Water)	45°C (pH 2.5)	KP1.2G; KP1.2Y
						45°C (pH 1.5)	
Rotokawa	RK1.1	E18788468 N5719957	Soil on embankment of stream, in fine sand and pumice	21.5	2.7-3.00	N/A	N/A
	RK1.2	E18788462 N5719955	Inside a slight trench of the embankment	31.5	1.7-1.9	45°C (pH 2.5)	RK1.2
						45°C (pH 1.5)	
	RK1.3	E1878463 N5719944	Edge of water stream in Silica rock formations	48	3.2	45°C (pH 2.5)	RK1.3

3.3.2 Isolation and Morphology

Cyanidiales isolates exhibited two distinct colony morphologies (MAll-Ht solid medium), either growth as a flat-continuous green mat (fig. 3.3 A) or growth as discrete raised yellow-green colonies (fig 3.3 B). Both colony morphologies were observed and isolated from many of the enrichments (fig. 3.3 C,D). Ten isolates were generated from the seven enrichments; four "green" isolates (CM1.1G, CM1.2G, CM1.3G, KP1.2G) and six "yellow" isolates (CM1.2Y, CM1.3Y, CM1.3.55, KP1.2Y, RK1.2, RK1.3).The colony morphologies were consistent with morphology observed under a microscope. All the yellow isolates had large round cells with endospores observed (figs. 3.4, 3.5) Whereas all green isolates were all smaller club-shaped cells and potential endospores were observed (figs. 3.6, 3.7).

Four isolates:CM1.3G, KP1.2G, CM1.3Y and RK1.3, were prioritised for further characterisation. This gave two green isolates and two yellow isolates that were geographically distant. Furthermore, preliminary data from the enrichments indicated that CM1.3Y isolate could survive up to 55°C and pH below 1.0.

One challenge was when the yellow isolates were grown on MAll-Ht plates, the isolates had a very long lag-phase (>21 days observed) when transferring to autotrophic conditions, which has been an issue noted with *G. sulphuraria* strains[91]. This is likely due to the reduced photosynthetic pigment content and the time required to replenish these pigments. This was the reason why the media for isolates RK1.3Y and CM1.3Y was supplemented with low concentrations of glucose in pH range (section 3.2.6) and nitrogen utilization (3.2.5) experiments.

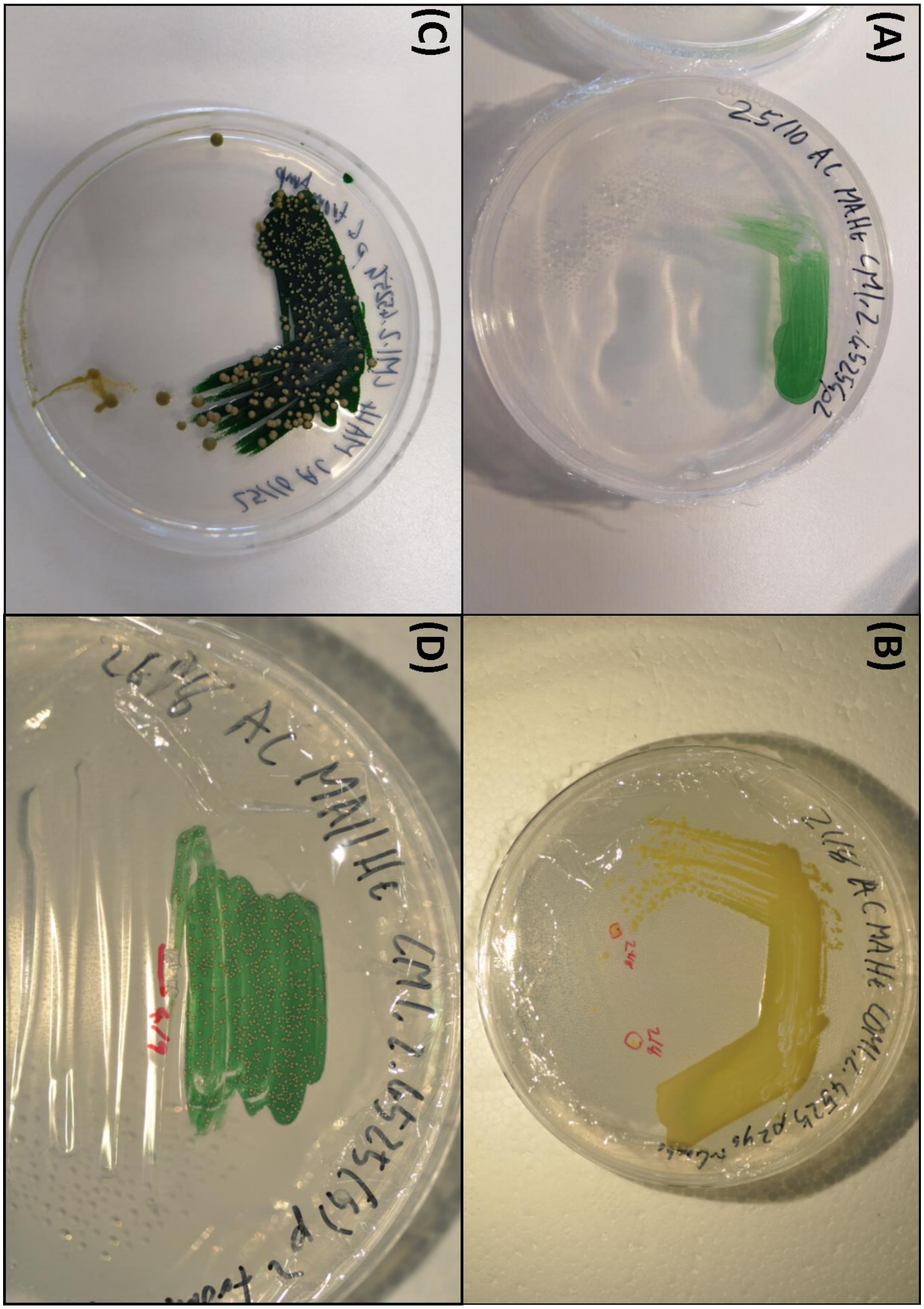


Figure 3.3: The different colony morphologies of Cyanidiales cells observed when cultured on Mall-Ht medium. All plates shown are from a single enrichment (CM1.2 at 45°C and pH 2.5). **A)** Is an axenic culture of CM1.2G a green isolate. **B)** Is an axenic culture of CM1.2Y a yellow isolate. **C,D)** Show mixture cultures during the isolation process. The difference of raise discrete colonies of the yellow isolate and the flat continuous mat of the green isolates are observable in C)

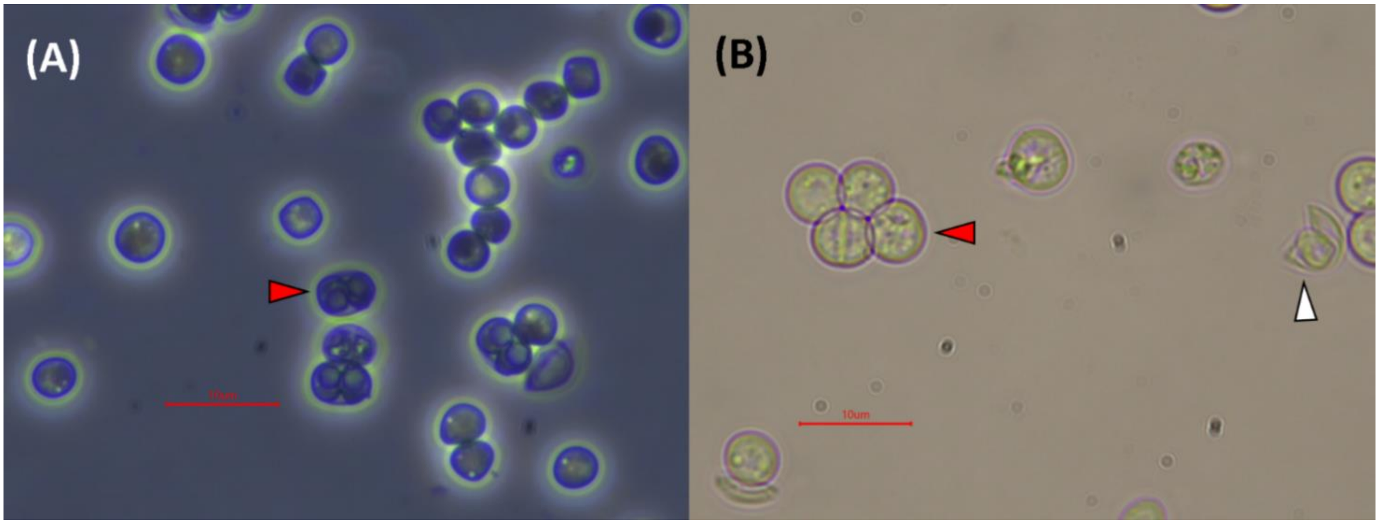


Figure 3.4: Microscope images from axenic cultures of the yellow isolate **CM1.3Y** as observed using phase-contrast(A) and bright field (B) with 100X oil immersion objective. Red triangles indicate vegetative cells with endospores. White triangles indicate cell wall debris from lysed vegetative cells. Scale bar is 10µm

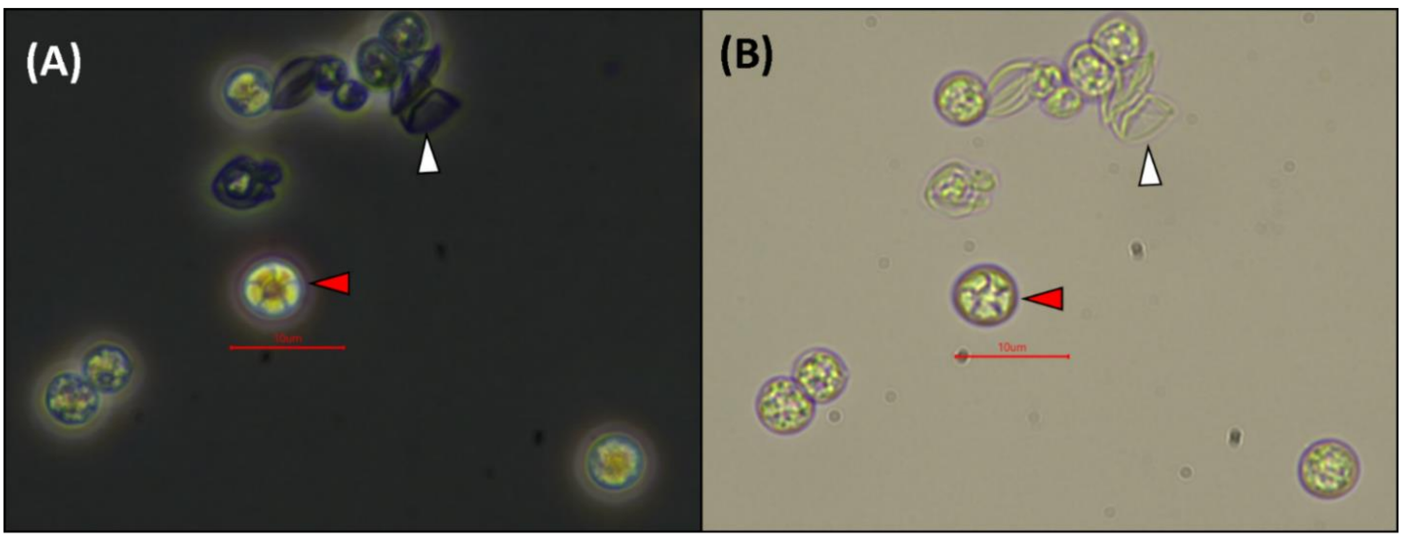


Figure 3.5: Microscope images from axenic cultures of the yellow isolate **RK1.3** as observed using phase-contrast(A) and bright field (B) with 100x Oil immersion objective. Red triangles indicate vegetative cells with endospores. White triangles indicate cell wall debris from lysed vegetative cells. Scale bar is 10µm

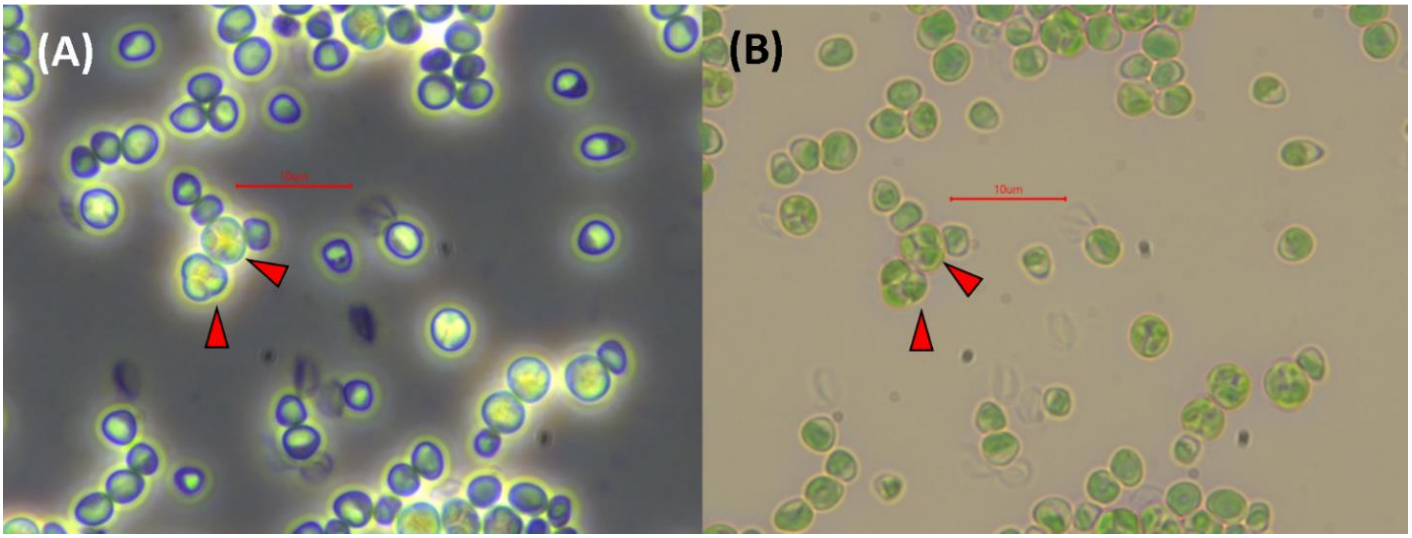


Figure 3.6: Microscope images from axenic cultures of the green isolate **CM1.3G** as observed using phase-contrast (A) and bright field (B) with 100x Oil immersion objective. Red triangles indicate vegetative cells with endospores. Scale bar is 10µm

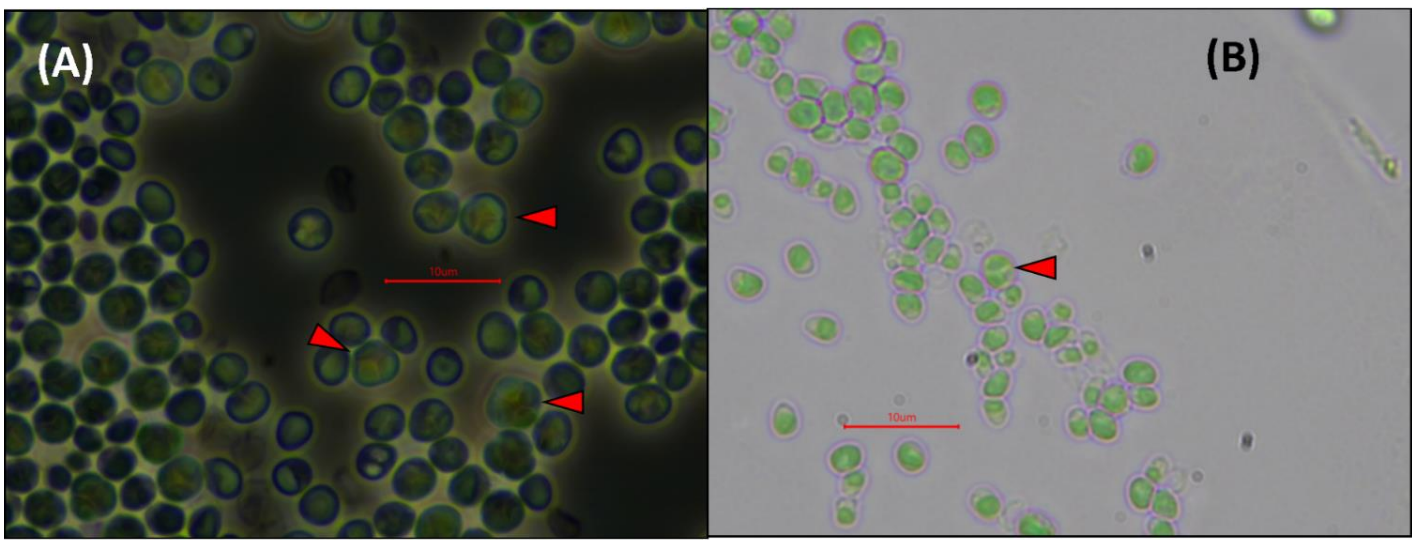


Figure 3.7: Microscope images from axenic cultures of the green isolate **KP1.2G** as observed using phase-contrast(A) and bright field (B) with 100x Oil immersion objective. Red triangles indicate vegetative cells with endospores. Scale bar is 10µm

3.3.3 Nitrogen utilization

Similar results were observed for both green isolates KP1.2G and CM1.3G. After one week, no growth was observed in the N-free serum bottles. Similar growth was observed in all the NH_4^+ and urea bottles (figs. 3.8, 3.9 table3.2). The growth observed in the nitrate bottles was fainter and less green (fig.3.8 table 3.2) than the NH_4^+ and urea bottles. After a three-week period, a single nitrate replicate for KP1.2G (fig 3.8 D), reached a similar cell density and colour to NH_4^+ bottles. Although, the observed growth was lower in the KP1.2G nitrate replicates (fig 3.8) and all of the CM1.3G nitrate replicates. This indicates that the isolates KP1.2G and CM1.3G can utilize ammonium, urea and nitrate as sole nitrogen sources. However, it appears nitrate is not an optimal nitrogen source compared to urea or ammonium.

For the yellow isolates, CM1.3Y and RK1.3, growth was observed only in the NH_4^+ replicates. So RK1.3 and CM1.3Y can utilize only ammonium as the sole nitrogen source.

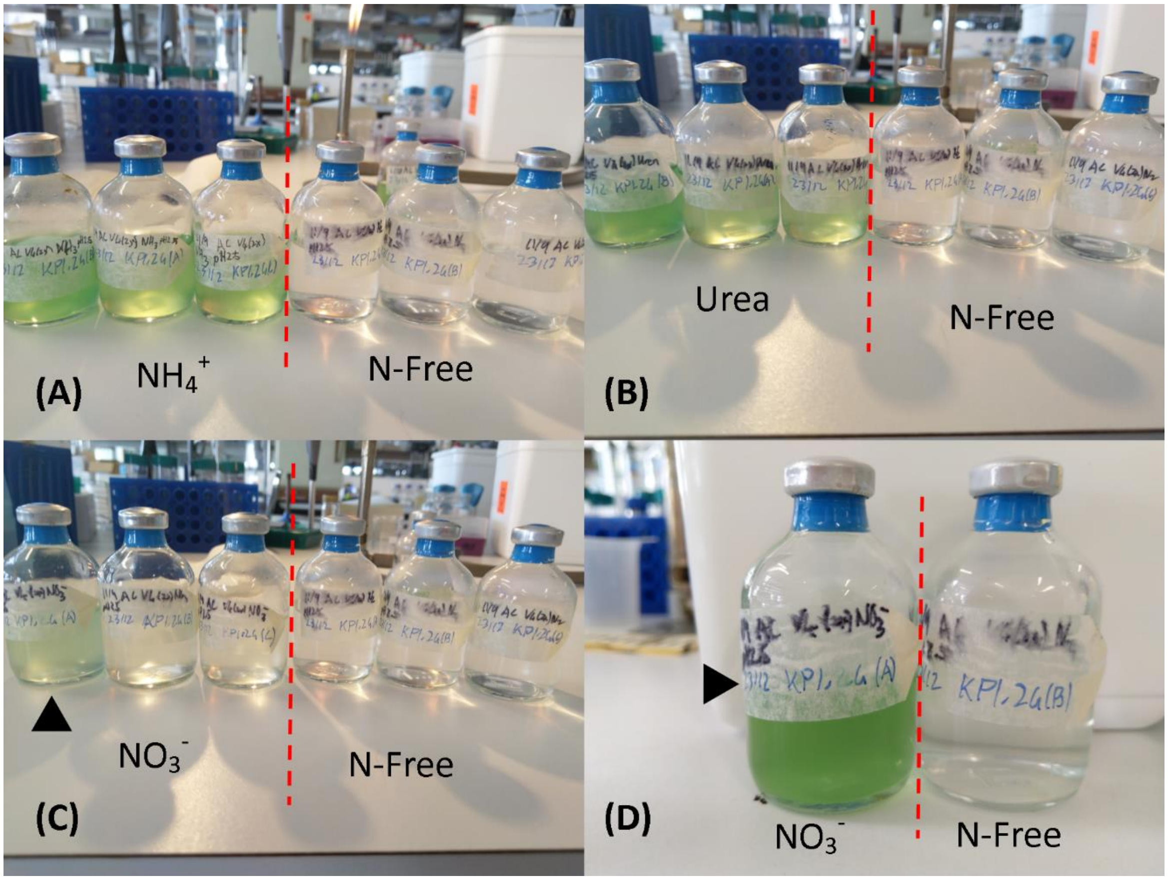


Figure 3.8: Growth of KP1.2G utilizing different Nitrogen sources as a representative of the green isolates. A-C) Is observed growth on the Nitrogen sources next to Nitrogen free media after 7 days A) Is growth with ammonium B) Is growth with Urea and C) Is growth with nitrate. D) Is a nitrate bottle after 21 days indicated by the triangle in C).

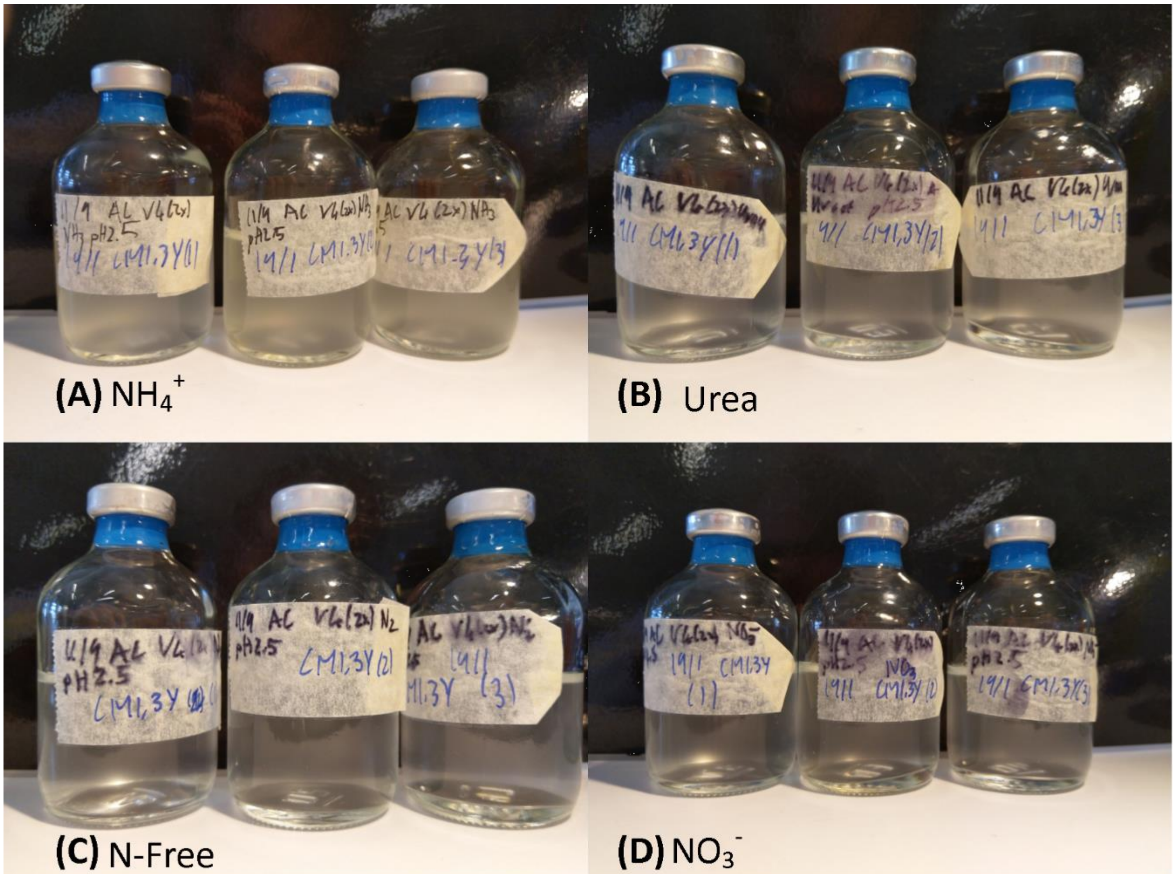


Figure 3.9: Growth after 7 days of CM1.3Y utilizing different Nitrogen sources as a representative of the yellow isolates. A) Is growth with ammonium B) Is growth with Urea and C) Is growth with nitrate. D) Is growth with nitrate.

Table 3.2. Relative growth of Cyanidiales isolates on. +++; strong growth, ++; growth, +; weak growth, +/-; variable growth, -; no growth

	Green isolates		Yellow isolates	
Nitrogen source	CM1.3G	KP1.2G	RK1.3	CM1.3Y
Nitrogen free medium	-	-	-	-
Ammonium	+++	++++	+++	+++
Urea	+++	+++	-	-
Nitrate	+	++	-	-

3.3.4 pH range

The growth of the Cyanidiales isolates were tested at different pHs (discrete values tested were; 0.0; 0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.0; 4.5; 5.0; 6.0;7.0). Each isolate was tested in triplicate and non-inoculated control were included (Table 3.3). For the green isolates, CM1.3G and KP1.2G, growth was observed from pH 1.0 through to pH 7.0 with no growth was observed at 0.0, 0.5 or 2.5 pH (fig 3.10. A, B). The lack of growth at pH 2.5 was unexpected, particularly as during routine cultivations both strains grew well at pH 2.5. Accordingly, it is anticipated that the bottles had erroneously not been inoculated. Based on the density and colour (fig.3.10 B; Table 3.3), the optimal pH of CM1.3G is approximately pH 2.0, with reduced growth observed at pH 1.5 and 1.0 (fig.3.10 B). Similar growth was observed at pH 4.0 to 7.0. With isolate KP1.2G, optimal growth occurs at pH 3.0-3.5 (fig3.10 A) and, as compared to CM1.3G, greater growth was observed at pH 1.0 and 1.5. Growth remained similar above pH 4.0.

For the yellow isolates, CM1.3Y and RK1.3, growth was observed from pH 0.5 through to pH 7.0 with no growth was observed at 0.0 pH (fig 3.10. C, D). For both CM1.3Y and RK1.3 it is difficult to determine based on the density and colour, as similar growth was observed at pH 1.5 to 7.0 (fig.3.10 C,D; Table 3.3). Growth reach saturation within the week due to increased growth rate of heterotrophic conditions and carbon limitation. Reduced growth observed at pH 1.0 and 0.5 (fig.3.10 C,D).

The pH of the uninoculated controls were stable after autoclaving and after the incubation period at 45°C, under constant illumination. The pH test did indicate that the isolates were able to acidify the medium at more neutral pH (>3.0). CM1.3G was acidified to pH 4.0, KP1.2G to pH 3.5-4.0 and the yellow isolates acidified media to pH 3.0-3.5. The pH of the more acidic inoculated bottles was unchanged.

Table 3.3 Relative growth of Cyanidiales isolates at different pHs. +++; strong growth, ++; growth, +; weak growth, +/-; variable growth, -; no growth

	Green isolates		Yellow isolates		Non-inoculated control
pH	CM1.3 G	KP1.2 G	RK1.3	CM1.3 Y	
0	-	-	-	-	-
0.5	-	-	+	+	-
1	+	++	++	++	-
1.5	+	++	+++	+++	-
2	+++	++	+++	+++	-
2.5*	-	-	+++	+++	-
3	+	+++	+++	+++	-
3.5	++	+++	+++	+++	-
4	++	++	+++	+++	-
4.5	++	++	+++	+++	-
5.0	++	++	+++	+++	-
6.0	++	++	+++	+++	-
7.0	++	++	+++	+++	-

*; no growth at pH 2.5 by xxx was likely due to non-inoculation.

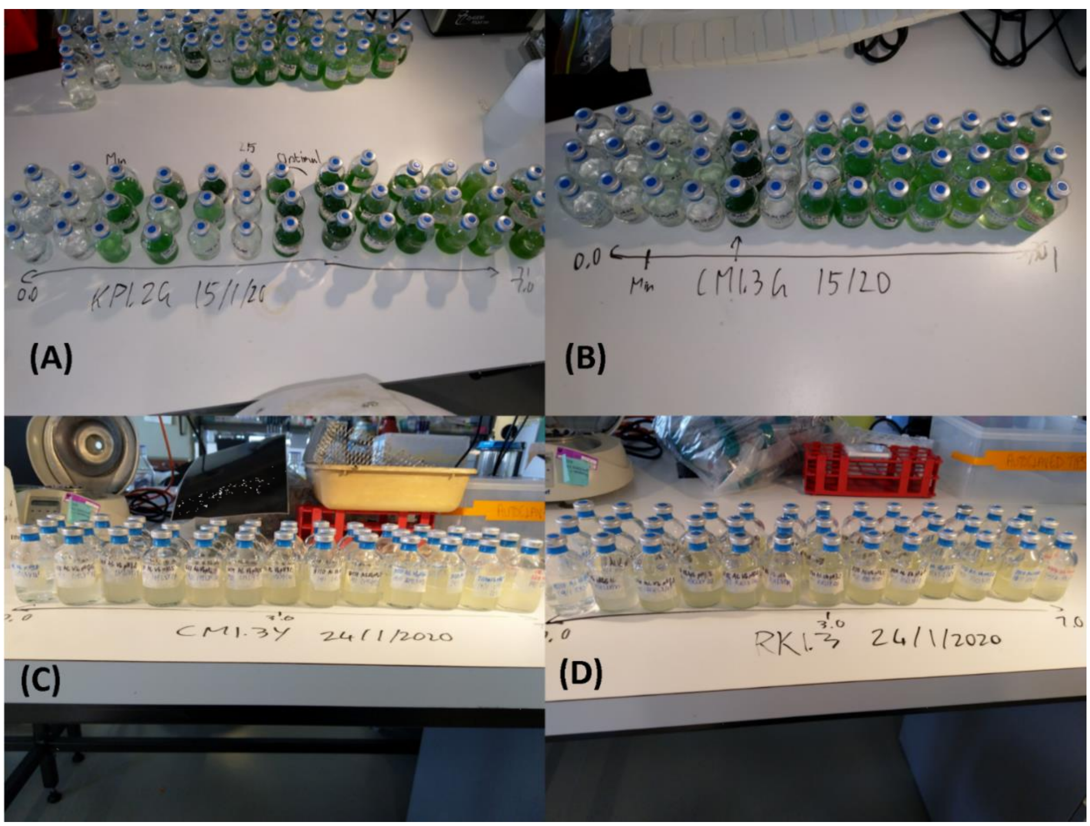


Figure 3.10: The pH growth range of A) KP1.2G B) CM1.3G C) CM1.3Y and D) RK1.3. Photo was taken 21 days after inoculation for A) and B). For C) and D) photo was taken 7 days after inoculation. Bottles are arranged left to right from most acidic (pH0.0) to least acidic (pH 7.0).

3.3.5 Temperature Range

For all isolates, growth was observed on plates incubated at 39°C, 45°C and 50°C. For the green isolates and RK1.3, the greatest growth was observed on the 45°C plates, while the greatest growth observed for CM1.3Y was on the 50°C plate. No growth was observed on the room temperature plates or the 54°C plates. (Fig3.11. table3.4)

Table 3.4 Relative growth of Cyanidiales isolates at different temperatures. +++; strong growth, ++; growth, +; weak growth, +/-; variable growth, -; no growth

	Green isolates		Yellow isolates	
Temperature (°C)	CM1.3G	KP1.2G	RK1.3	CM1.3Y
~25 (Room temp)	-	-	-	-
39	++	++	++	+
45	+++	+++	+++	++
50	+	+	++	+++
55	Not tested	Not tested	-	-
60	Not tested	Not tested	-	-

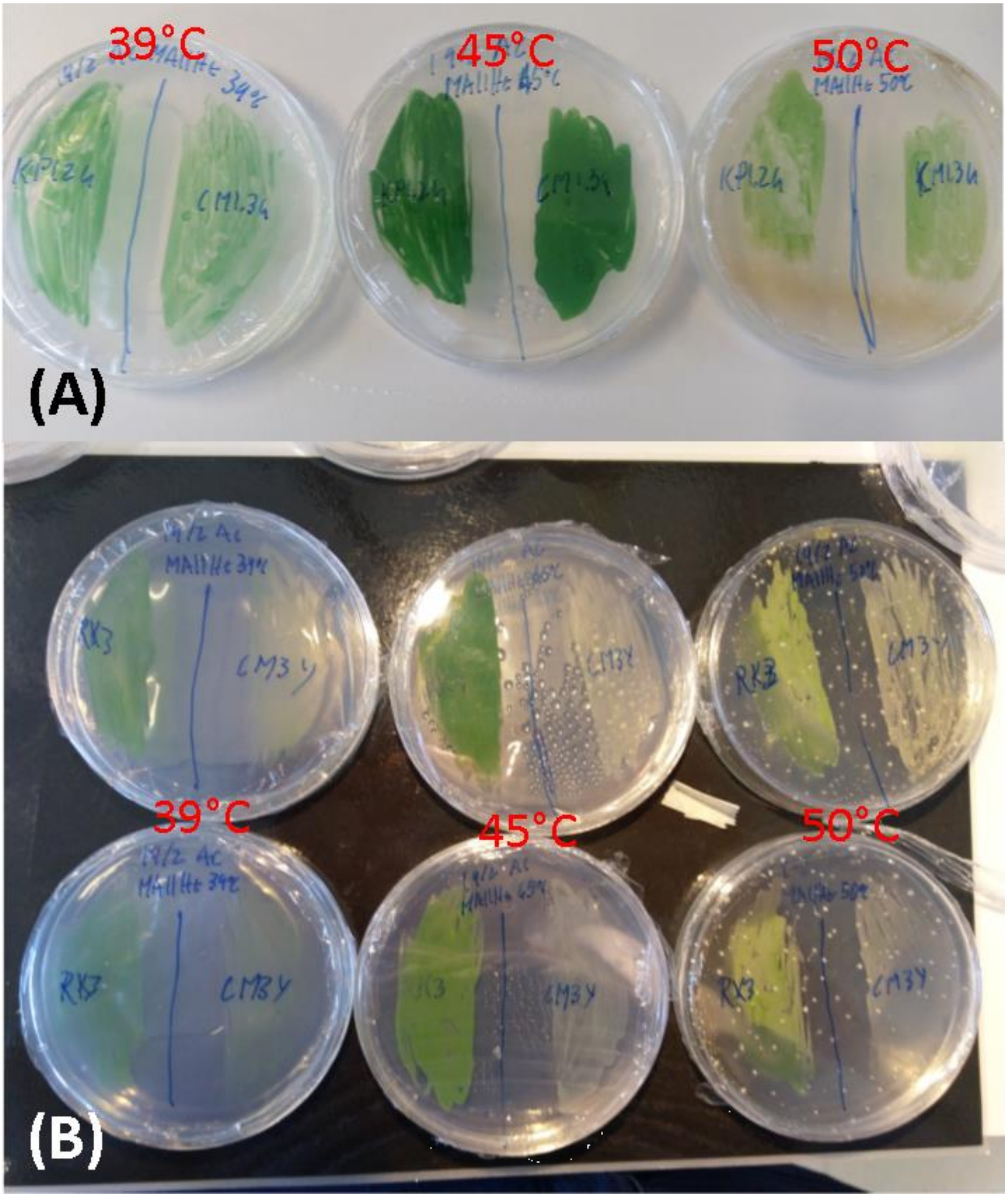


Figure 3.11: Growth temperature range of isolates, images of plates incubated at 39°C, 45°C and 55°C. A) Growth green isolates KP1.2G, on left side of plates, and CM1.3G on the right side of the plates. B) Growth yellow isolates RK1.3 on left side of plates, and CM1.3Y on the right side of the plates.

3.3.6 Heterotrophic growth on glucose or glycerol

No growth of the green isolates was observed on either carbon source in the dark. Growth of the yellow isolates was observed on both carbon sources, but growth on glycerol was less.

3.3.7 Fatty acid profile

Table 3.5 FAME profile of isolates

	KP1.2G	**CM1.3G**	**CM1.3Y**	**RK1.3Y**
	%	**%**	**%**	**%**
14:0	1.3	1.0	0.2	0.3
14:1(n-5)	0.5	0.4		
15:0	1.0	1.0	0.3	0.4
16:0	35.4	36.2	32.5	33.2
16:1(n-7)	0.2	0.1	0.4	0.5
16:1(n-5)	0.2	0.3	n.d.	n.d.
16:2	1.5	1.4	0.8	0.9
17:0	1.1	1.1	0.2	0.2
17:1	0.5	0.2	n.d.	n.d.
18:0	6.3	7.5	4.0	3.3
18:1(n-9)	22.3	21.2	26.6	22.7
18:2(n-6)	26.9	27.3	27.0	29.7
18:3(n-3)	n.d.	n.d.	5.8	6.4
20:1	0.3	0.2	0.2	0.3
20:2	0.4	0.5	0.6	0.6

Results are for fatty acids present above 0.2% (of total fatty acids) in at least one sample. The predominant single species of fatty acid for all isolates was palmitic acid (C16:0). At 35 to 36%, for KP1.2G and CM1.3G, and at 32-33% for CM1.3Y and RK1.3. All isolates have high concentrations of the C18 unsaturated fatty acids (as oleic and linoleic acid), but α-linolenic acid was only detected in isolates RK1.3 and CM1.3Y.

3.3.8 TLC- Polar lipids

Total cellular lipids were extracted from isolates CM1.3G, CM1.3Y, KP1.2G and RK1.3. A 1D-TLC gel was run to compare the polar lipid profiles of all the isolates. As 1D-TLC lacks resolution the polar lipids profiles were confirmed running 2D-TLC gels of individual isolates. Glycolipid, phospholipids and total lipids were highlighted using different staining methods (section 3.2.10). The phospholipids phosphatidylinositol (PI) and phosphatidyl-glycerol (PG) can have similar profile as other phospholipids on TLC gels, for example, the PI spots (figs. 3.13-3.16) could be LysoPC (Phosphatidylcholine intermediate with a single fatty acid). Therefore, these phospholipids will be confirmed with LS-MS but at the time of writing the results are not available. The PG and PI spots were marked with asterisks to highlight this.

For the results the cellular lipids of all isolates (fig. 3.12 A) contain the glycolipids MGDG DGDG (mono and digalactosyly-diacylglycerol respectively) and SQDG (sulfoquinovosyldiacylglycerol). All isolates also contain the phospholipids PE (Phosphatidylethanolamine) PC(Phosphatidylcholine) and PI* (figs 3.12-3.16. B). The phospholipid PG* was detect in the green isolates CM1.3G (fig. 3.14 B) and KP1.2G (fig. 3.13 B) but not in the yellow isolates CM1.3Y (fig. 3.15 B) and RK1.3 (fig. 3.16 B). When comparing the polar lipid profiles of the isolates (fig. 3.12), the dominate polar lipids of the green isolates are the glycolipids MGDG and DGDG. Whereas the dominate polar lipids of the yellow isolates are the phospholipids PE and PC (fig. 3.12)

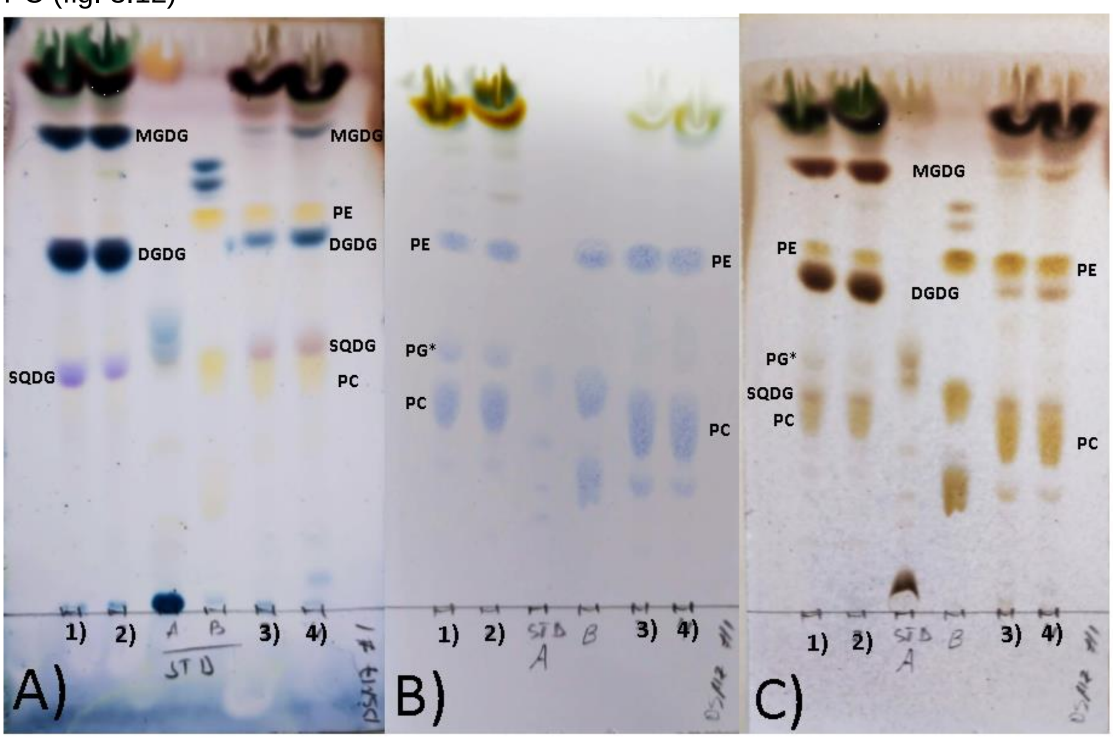


Figure 3.12: 1D-TLC results, the solvent system is chloroform:methanol:water (65:25:4, by volume) in the up direction. For the columns 1) KP1.2G 2) CM1.3G, then 2 columns of known standards, 3) CM1.3Y 4) RK1.3. A) Anthrone stain, glycolipids appear blue and phospholipids appear yellow. B) Molybdate stain highlights only phospholipids C) Charred stain highlights all lipids. Lipids labelled with asterisks (*) need to be confirmed by Mass-spectrometry analysis

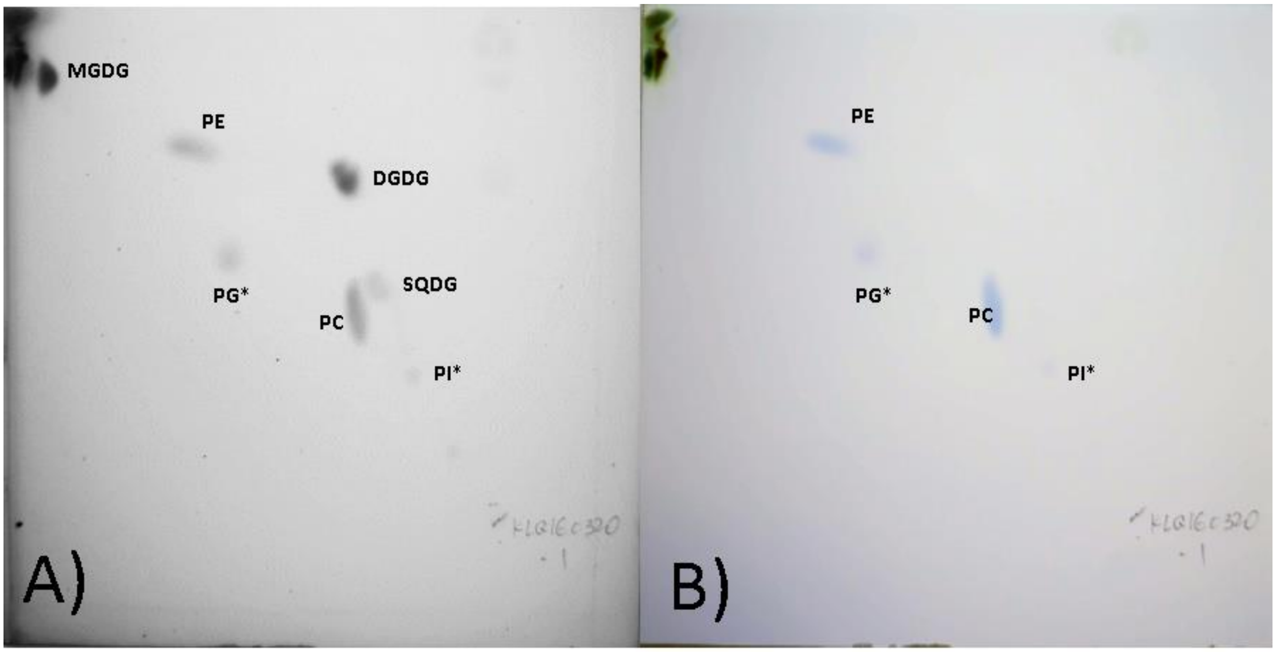


Figure 3.13: 2D-TLC results from green isolate **KP1.2G**. For the solvent system first direction (up) is developed in chloroform:methanol:water (65:25:4, by volume), and the second (left) in chloroform:methanol:acetic acid:water (80:12:15:4, by volume). A) Charred stain highlighting all lipids. B) Molybdate stain highlighting only phospholipids. Lipids labelled with asterisks (*) need to be confirmed by Mass-spectrometry analysis.

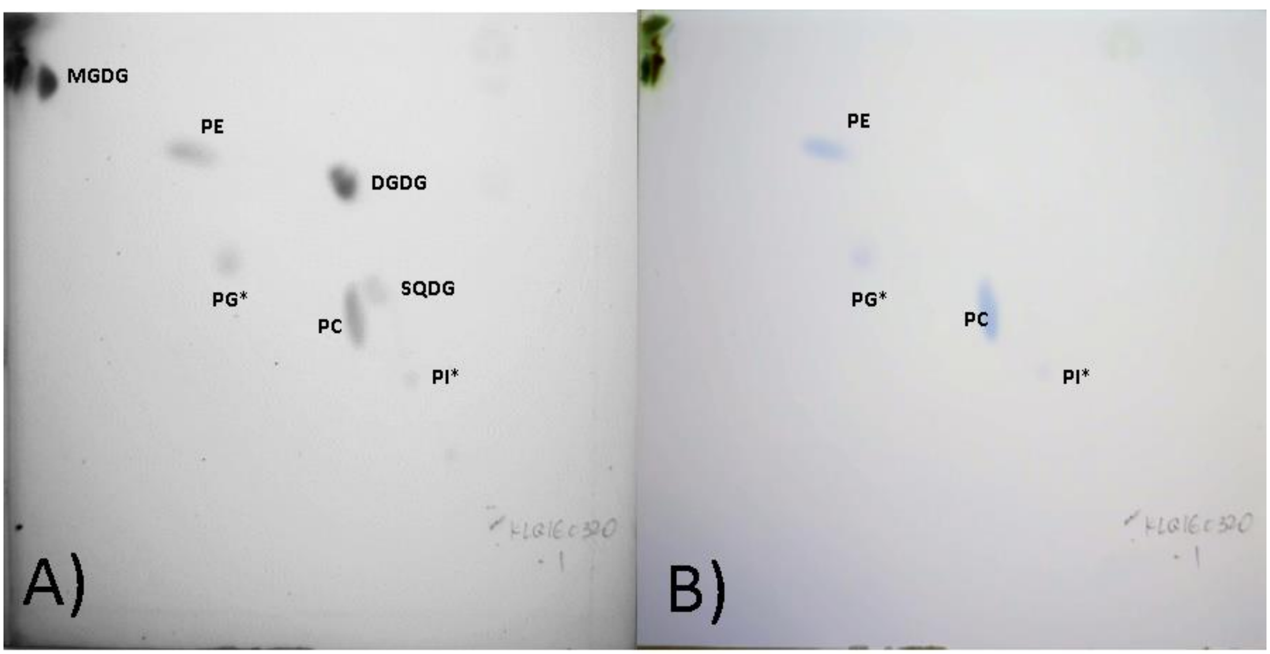


Figure 3.14: 2D-TLC results from green isolate **CM1.3G**. For the solvent system first direction (up) is developed in chloroform:methanol:water (65:25:4, by volume), and the second (left) in chloroform:methanol:acetic acid:water (80:12:15:4, by volume). A) Charred stain highlighting all lipids. B) Molybdate stain highlighting only phospholipids. Lipids labelled with asterisks (*) need to be confirmed by Mass-spectrometry analysis.

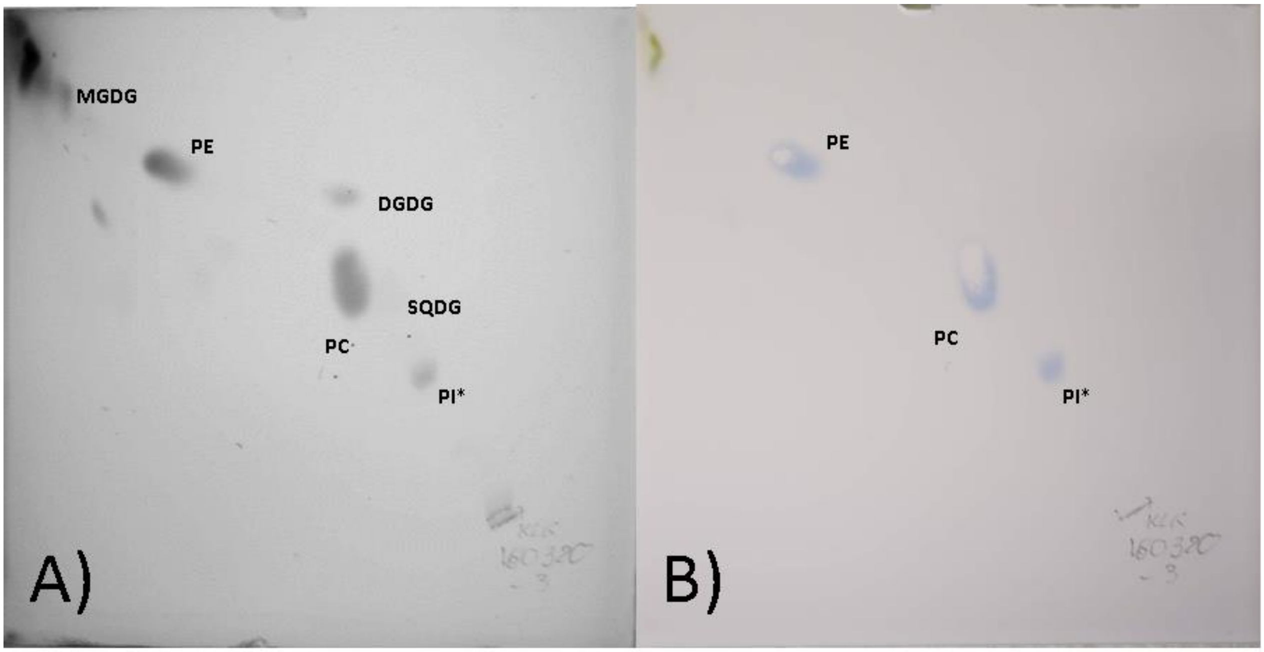


Figure 3.15: 2D-TLC results from yellow isolate **CM1.3Y**. For the solvent system first direction (up) is developed in chloroform:methanol:water (65:25:4, by volume), and the second (left) in chloroform:methanol:acetic acid:water (80:12:15:4, by volume). A) Charred stain highlighting all lipids. B) Molybdate stain highlighting only phospholipids. Lipids labelled with asterisks (*) need to be confirmed by Mass-spectrometry analysis.

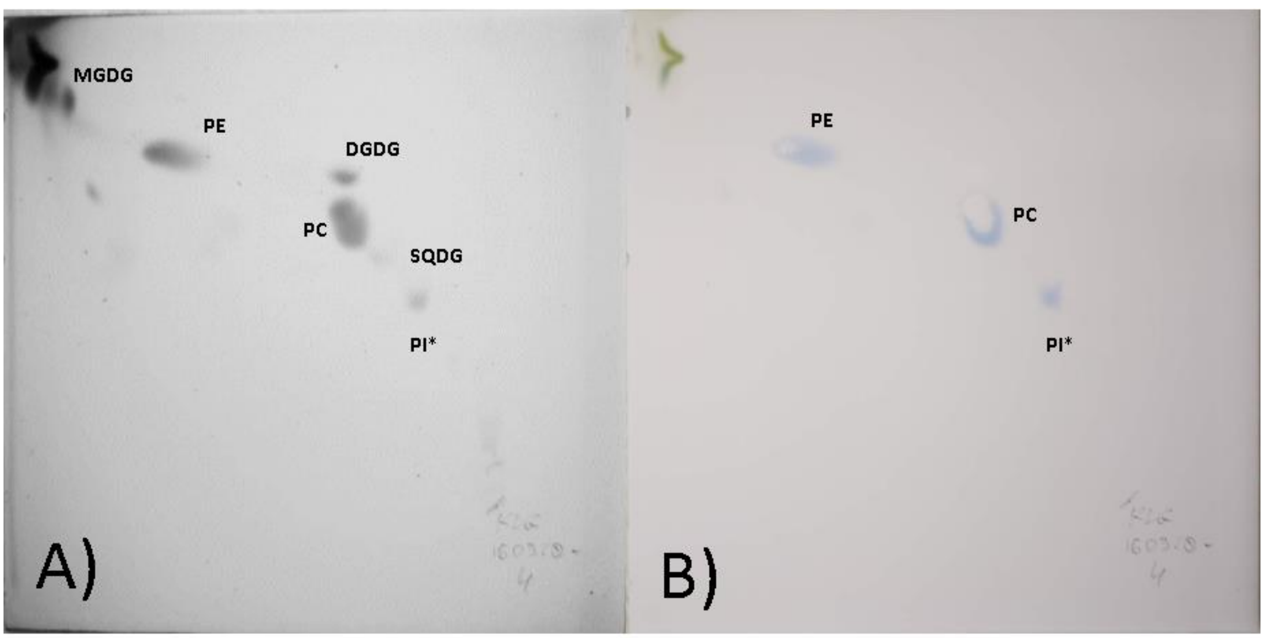


Figure 3.16: 2D-TLC results from yellow isolate **RK1.3**. For the solvent system first direction (up) is developed in chloroform:methanol:water (65:25:4, by volume), and the second (left) in chloroform:methanol:acetic acid:water (80:12:15:4, by volume). A) Charred stain highlighting all lipids. B) Molybdate stain highlighting only phospholipids. Lipids labelled with asterisks (*) need to be confirmed by Mass-spectrometry analysis.

3.3.9 Summary of results

Table 3.6 Summary of the results for phenotypic characterisation of Isolates

Characteristic	CM1.3G	KP1.2G	CM1.3Y	RK1.3
Colony morphology on MAll-Ht	Green flat-mat	Green Flat-mat	Raised discrete colonies- Yellow (young) Yellow-green (Old)	Raised discrete colonies- Yellow (young) Yellow-green (Old)
Cell morphology	Small (2-4µm) oblong juvenile cells. More round (4-5µm) vegetative cells.	Small (2-4µm) oblong juvenile cells. More round (4-5µm) vegetative cells	Large round cells (~4-6µm). Vegetative cells can reach up to 10µm	Large round cells (~4-6µm). Vegetative cells can reach up to 10µm
Endospore	Yes	Yes	Yes	Yes
Heterotrophic growth	No	No	Yes	Yes
Nitrogen sources	NH_4^+, NO_3^-, Urea	NH_4^+, NO_3^-, Urea	NH_4^+	NH_4^+
pH for Growth	1.0-4.0+ (*)	1.0-3.5+ (*)	0.5-3.0 (*)	0.5- 3.0 (*)
Optimal pH	2.0-2.5	2.5-3.5	N.d.	N.d.
Temperature range (measured)	39-50°C Optimal= 45°C	39-50°C Optimal= 45°C	39-55°C* Optimal= 45°C	39-50°C Optimal= 45°C
α-linolenic acid	Absent	Absent	Present	Present
Major Polar lipid	Glycolipids MGDG, DGDG	Glycolipids MGDG, DGDG	Phospholipids PE, PC	Phospholipids PE, PC

3.4 Discussion

3.4.1 Isolation and Enhanced growth on Medium with Glycerol and Glucose

Isolation of axenic cultures was the most time-consuming process of this Master's thesis. From having successful enrichments at the start of May 2019, it took until the start of October 2019 to have multiple cultures that I had confidence were free of bacterial contamination and had a single Cyanidiales morphology(section 3.3.2). Multiple other methods and media were attempted concurrently before the development of the final MAll-Ht medium and method. Serial dilution to extinction method was attempted, however it took over two weeks to detect growth, even under a microscope, between passages. Additionally, bacterial contamination was persistent as the bacterial cell count was magnitudes higher than the Cyanidiales cell count. Isolation on other solid media was attempted as well. Initially, 2xV4 medium [23,135] plates were attempted, given growth in liquid V4 medium had been observed in the enrichments. The 2xV4 plates did not support any growth. Next, modified Allen's medium (MAll [136]) was attempted, a recipe from the literature which is derived from the original Allen's medium [48] and optimised for the growth of *C. merolae* 10D [136]. There was limited success, growth was observed but the rate was low and there was persistent bacterial contamination. The enrichment of CM1.3 at 55°C and pH 2.5, and 45°C at pH 0.7, did grow discrete pickable colonies without bacterial contamination.It is likely that the extreme conditions of the growth medium, restricted the growth of the contaminating bacteria reducing competition and allowing for better growth on the MAll plates. Other method to reduce bacterial contamination were also tested. Different exposure durations to UV light was tested, however any duration that was effective at reducing bacterial growth completely inhibited Cyanidiales growth. Several antibiotics were tested independently, including others that were not used in the final isolation. For example, tetracycline was effective against bacterial contamination, but it also bleached the Cyanidiales cells. Antibiotics such as ampicillin and vancomycin were known not to have a significant effect on *C. merolae* growth [136], however independently did not completely inhibit bacterial growth. Hence the use of the ampicillin, vancomycin and erythromycin in combination, which was effective against the bacterial contamination, had no observable effect on Cyanidiales growth. The Cyanidiales growth on standard MAll was still slow, meaning the that the above mentioned hypothesis that competition from bacterial contamination was the primary limiting factor, was incorrect. Heterotrophic media has been used to isolate Cyanidiales [61,91] before, but there were two main concerns; increased bacterial growth and biased selection towards facultative heterotrophs like

G. sulphuraria. As an effective antibiotic treatment had been developed, the concern about bacterial contamination had been addressed. Furthermore, as *G. sulphuraria* was one of the species targeted for isolation, the bias for selection of this species was not a concern. A heterotrophic modified Allen's medium (MAll-Ht) plates was developed by addition of $3gL^{-1}$ glucose and $3gL^{-1}$ glycerol. As expected, without antibiotic treatment, bacteria out competed the Cyanidiales, but could be reduced significantly with the antibiotic treatment.

What was unexpected was the enhanced growth of both the green and yellow isolates, when incubated at 45°C under constant illumination.

Variation in pigmentation, when grown using sugars as the carbon source, are an established phenotypes of *G. sulphuraria* and has been used previously to isolate individual *G. sulphuraria* strains from each other[91]. There is evidence that the yellow isolates are *G. sulphuraria* that are simply metabolizing the glucose for a higher growth rate. However, there is also evidence that the green cultures are not *G. sulphuraria* (section 3.4.4) and no heterotrophic growth of KP1.2G and CM1.3G has been observed on glucose or glycerol in the dark. A small percentage of glycerol (0.4%) in Modified Allen's medium has been observed to enhance the growth of *C. merolae* 10D [136], a well studied obligate photoautotroph [30,31,57].

I have three hypotheses as to why the sole modification of the addition of glycerol and glucose enhanced the growth of autotrophic green isolates:

1) Enhanced growth is due to plates having lower pH and therefore conditions are more optimal for growth. The final pH of the sterile MAll-Ht plates was measured to be lower than the final pH of either the MAll or 2xV4 plates (pH 4 to 4.5 and 5 respectively), despite all being adjusted to pH 2.5 prior to autoclaving. Meaning that either the sugars are acting as buffers or have oxidized into organic acids, however I have no plausible evidence how this could occur mechanically.
2) The mixture contains glycerol and sulphuric acid at pH 2.5 which is heated to 121°C in the autoclave. There is the potential that organosulfates are being produced[140]. Organosulfates include common surfactants, such as sodium lauryl sulfate (SLS), which could reduce the water tension and hence reduce desiccation or osmotic stress.
3) A comparative genome study of *G. sulphuraria* 074W and *C. merolae* 10D, indicated that both had intact carbon metabolic pathways, including key enzymes like glycerol kinases[57]. However, *the C. merolae* genome lacked many carbohydrate transporters found in the *G. sulphuraria* genome, like glycerol permeases[57]. This means that *C. merolae* can metabolise carbohydrates but lacks the mechanisms to take up carbohydrates from the environment. Therefore, it is possible that the glycerol and

glucose are entering the cell of the autotrophic green isolates through passive non-specific/non-selective channels and are a supplemental energy and/or carbon supply. It is likely that these carbohydrates do not enter the cell at a rate sufficient enough to be the sole energy or carbon source.

I considered the third hypothesis the most probable, however no experiments were done to test these hypotheses due to time constraints. A quick experiment would be to add the glycerol/glucose as a filter sterilised solution to the medium, post autoclaving. There will be no chemical change to the glycerol or glucose, so if there is enhanced growth, it is acting as a supplemental energy/carbon supply, supporting the third hypothesis. If growth is not enhanced, a small percentage of organosulfate surfactant could be added post autoclaving to test the second hypothesis. The first hypothesis could be tested by generating more acidic plates, however there is evidence that the isolates can acidify media (section 3.3.4 & 3.4.2). Therefore, if this was the primary factor, a longer-lag followed by normal growth would be expected. Additionally, the difference between pH 4.0 and pH 4.5 should not be that significant.

In lieu of the mechanical reason for the enhanced growth of green isolates, the growth and isolation, on MAll-Ht medium plates, allowed for the detection of multiple strains within a single sample, made selection and isolation easier due to the contrast in colours, and resulted in higher growth rate, therefore less time between passages. There were issues as there is a long-lag phase for the yellow isolates when switching between the plates and autotrophic media and the added glucose does promote bacterial growth. The bacterial growth can be dealt with using antibiotics and correct aseptic techniques.

While the growth of Cyanidiales that lack a cell wall, such as *C. merolae,* has not been observed on MAll-HT in this thesis, this is a novel modification and is recommended for the isolation of all future Cyanidiales.

3.4.2 Acidification of media and Ammonium uptake.

Experiments designed to determine the pH ranges and optima showed that the Cyanidiales isolate were capable of acidifying the media to pH values as low as 3.0. The acidification was similar between the green isolates, CM1.3G and KP1.2G, grown in autotrophic conditions and the yellow isolates, RK1.3 and CM1.3Y, grown in mixotrophic conditions. This has been observed in Cyanidiales species before including *C. caldarium* [141], *G. maxima* (peer-review pending [142]) and NZ strains [143]. Lowell and Castenholz 2013 hypothesise that Cyanidiales acidify the environment with ATP-dependent H^+ efflux and it is an adaption to modify the pH of the environment for increased growth and survivability. However, there is a well-documented relationship between the uptake of ammonium and the decrease of external pH, documented in *C. caldarium* [141] and other photosynthetic eukaryotes including green algae and land plants [144–146]. This acidification is a byproduct of the assimilation of ammonium. While there are many pathways, the amination of organic compounds generates excess protons (equation 1) that need to be pumped out of the cytosol to maintain the near neutral internal pH [141,146,147].

Equation 1: $NH_4^+ + R = R\text{-}NH_3 + H^+$

The same phenomena are observed in non-acidophilic plants[144, 145, 147, 148], therefore acidification is a by-product of an essential pathway, not a mechanism to alter the pH of the environment. Also, the acidity of geothermal sites is most dependant on chemistry of the mineral and gas present[34, 87, 134]. The concentrations of this chemicals are high enough that the growth of micro-organism will have negligible effect on the pH of these environments. Given this I disagree with the hypothesis that the acidification capability of Cyanidiales strains has a significant effect on the selection or survivability of that strain in natural environments [142,143].
However, it can be exploited in batch bioreactors. A *G. sulphuraria* strain was able to reduce the pH of a 700L bioreactor to pH 4.0 [123], so there is the potential to not require an initial pH treatment of the media, at the cost of a longer lag-phase.

I was unaware of this acidification process prior to the pH experiments and, due to time constraints, I was unable to repeat the experiment. The V4 medium does contain phosphate as a buffer, however the total moles of phosphate is lower than the moles of ammonium (as NH_4Cl). To ensure the pH of the medium remained stable during growth, the concentration of buffers would need to be increased by magnitudes. The experiment did determine the pH minimum for all isolates tested and gave an approximation of optimal pH for the isolates CM1.3G and KP1.2G.

3.4.3 Nitrogen Sources

For isolates CM1.3Y and RK1.3, growth was only observed on NH_4^+ as the sole nitrogen source and is evidence that these strains are *G. sulphuraria*[31]. The isolates KP1.2G and CM1.3G could utilise a wide range of the tested nitrogen sources and growth was observed with NH_4^+, NO_3^-

and urea as the sole nitrogen sources. This does not provide strong evidence to determine taxonomy[31,36] but has other implications for biotechnology and HGT events in Cyanidiales[45]. The uptake and assimilation of nitrate, like ammonium, is linked to a change in external pH, however unlike ammonium, the external pH decreases with the uptake and assimilation of nitrate[141,148]. The nitrate is co-transported with a proton to balance charge [148,149] and the assimilation of nitrate requires the reduction of nitrate to ammonium (Equation 2), which requires the uptake of additional protons [141,148].

Equation 2: $NO_3^- + 8\ H_{(intracellular)} + 2\ H^+_{(extracellular)} \Rightarrow NH_4^+ + 3\ H_2O$

The process reduces the number of H^+ in the environment and is observed as an increase in the external pH. Given that most Cyanidiales are obligate acidophiles and the evidence that the V4 medium is not sufficiently buffered, I hypothesise that the poor growth observed on nitrate as the sole source of nitrogen was due to the alkalinisation of the medium. No pH measurements were taken, but the hypothesis could be tested by repeating the experiment with a higher concentration of buffers. This also has the potential to be exploited in bioreactors, as the balancing of nitrate and ammonium in the media could allow for metabolic pH control and prevent pH changes [144,145]. This will require empirical data for the specific strain used to implement effectively, as the uptake rate of nitrate or ammonium differs and will depend on the species. However, this could be key in choosing strains for biotechnological applications, as this is not possible with the RK1.3 and CM1.3Y since there is no evidence that these isolates can utilize nitrate.

The ability of KP1.2G and CM1.3G to utilize urea as the sole nitrogen source is a key observation, as almost all Cyanidiales cannot utilize urea due to genome reduction in the Cyanidiales common ancestor [76]. The ability to use urea requires the enzyme urease which is coded by 7 *ure* genes. Evidence from whole genome data indicates that *C. merolae* 10D has lost all of the seven *ure* genes[42,43,45], while *G. sulphuraria* 074W has only one *ure* gene[44,45]. The draft genome of *G.phlegrea* DBV009, an Italian strain that can utilize urea, has all seven *ure* genes, all of which have Eubacterial origins acquired through HGT [45]. The recent survey of 10 novel Cyanidiales draft genomes confirmed the findings and was unique to DBV009 [46]. The evidence indicates that the KP1.2G and CM1.3G isolates are not *G.phlegrea*, but rather belong to the distant Cyanidiophyceae family. This means that these isolates have gained the ability to use urea and have acquired a set of *ure* genes independently of *G.phlegrea* DBV009, in a separate HGT event(s) and likely a different origin. This would require the whole genome to be sequenced as it would be difficult to target the unknown *ure* genes with primers, but would be a novel discovery within this lineage.

3.4.4 Lipid profiles

The lipid profile of the Cyanidiales is atypical among Rhodophyta and photosynthetic eukaryotes. In terms of the fatty acids, no α-linolenic acid (C18:3(n-3)) has been detected in *Cyanidioschyzon* spp. or *Cyanidium* spp. [30,31,150]. This has not been observed in any other photosynthetic eukaryote and is rare in Cyanobacteria [30,150]. Major desaturase enzymes involved in the polyunsaturated fatty acid (PUFA) production pathways were not found in the *C. merolae* genome [150]. *Galdieria* spp. can produce α-linolenic acid, however the quantity is low in comparison to other algae[151], typically under 10% of total fatty acids even under optimal conditions [12,69–71]. Rhodophyta are characterised as having high levels of long-chain (C20) PUFAs like arachidonic acid (C20:4) and eicosapentaenoic acid (C20:5)[151,152]. However, these fatty acids had not been detected in Cyanidiales[150,153,154] including *Galdieria*[12,69,70], until recently when trace amounts of arachidonic acid and eicosapentaenoic acid was detected in a yet unidentified *Galdieria* strain[71]. Interestingly, the preliminary work on a draft genome sequence from the *Galdieria* is yet to identify any saturase for these long-chain PUFAs, therefore this indicates an alternative pathway[71]. The FAME profiles of the four isolates(table 3.5) is consistent with the results from Cyanidiales within the literature[12,70,71,150], with palmitic acid (C16:0) as the predominant fatty acid, followed by linoleic acid (C18:2(n-6)) and oleic acid (C18:1(n-9)). Although α-linolenic acid was not detected in either KP1.2G or CM1.3G isolates while is present in CM1.3Y and RK1.3 isolates at around 6% of the total fatty acids. This is strong evidence that the yellow isolates(CM1.3Y, RK1.3) are *Galdieria* while the green isolates(KP1.2G, CM1.3G) are members of the Cyanidiaceae family.

In terms of polar lipids, glycolipids such as mono and digalactosyly-diacylglycerol (MGDG, DGDG), and sulfolipids such as sulfoquinovosyldiacylglycerol (SQDG), are plentiful in plants and algae [150,152]. Phospholipids are also common among plants and algae, however some phospholipids like Phosphatidylcholine (PC) are not ubiquitous and are absent from some algae[155]. For the results, the glycolipids MGDG, DGDG, SQDG, and the phospholipids Phosphatidylethanolamine (PE) and PC, are present in all isolates. However, the predominant polar lipids for the green isolates (CM1.3G, KP1.2G) are glycolipids (MGDG, DGDG, figs 3.12, 3.13, 3.14), while the predominant polar lipids of the yellow isolates(CM1.3Y, RK1.3) are the phospholipids (PE,PC, figs 3.12, 3.15, 3.16). It can be difficult to accurately determine the Cyanidiales genera based on the polar lipid profile, as most data for *C. caldarium*[153,154] was generated prior to the revision of Cyanidiales genera [30] and hence it is difficult to determine if strains used are *Cyanidium* or *Galdieria*. However, there is modern empirical and genomic data for *C. merolae* [150,156]and *G. sulphuraria*[69,71],for example the major polar lipids for *C. merolae* are glycolipids such as MGDG and DGDG[150,157] while the major polar lipids for *G. sulphuraria* are phospholipids (PC,PE)[69,71]. There is evidence that this difference is due to a lack of polar lipid pathways and phosphate starvation. There are a number of metabolic pathways that plants and algae can use to produce PC, *C. merolae* can only produce PC through methylation of PE which is typically considered a minor pathway[155]. Additionally, none of the major betaine lipids have been detected in either *C. caldarium* [158]or *C. merolae* [150], but have been detected in some *G. sulphuraria* strains[69] however not all[71]. Betaine lipids are used by algae as a surrogate for phospholipids in phosphate limited conditions[152]. Phosphate limited conditions occur in an acidic environment, due to undissociated phosphate (H_3PO_4) becoming the predominant phosphate species as the pH decreases[69]. Undissociated phosphate is less bioavailable than the ionic phosphate species e.g. $H_2PO_4^-$[69]. There is evidence that *G. sulphuraria* is highly adapted to this phosphate limited environment, as phosphate uptake is highest in acidic conditions[110] and can excrete catabolic enzymes like acid phosphatase, which removes phosphate from exogenous organic compounds[44]. However, despite these adaptations in *G. sulphuraria* that have betaine lipids, the more acidic the growth conditions the greater the percentage of betaine lipids in relation to polar lipids[69]. This is evidence that phosphate starvation can still occur in *G. sulphuraria*. None of these adaptations to overcome phosphate limiting conditions, nor betaine lipids, have been observed in the Cyanidiaceae family[150,153,158]. Rather *C. merolae*, like plants, uses glycolipids as a surrogate for phospholipids under phosphate limiting conditions[152]. I would hypothesize that inhabiting acidic geothermal sites meant that phosphate limitation was a primary driving factor in the evolution of Cyanidiales. In the Cyanidiaceae family, the predominance of “primitive” Cyanobacteria pathways for polar lipids (i.e. the glycolipids) lowers the phosphate needs of the cells. Therefore, many of the eukaryotic phospholipid pathways could be lost to mutation and genome reduction[76], without adversely affecting the survival of organisms. In the Galdieriaceae family, adaptations overcame the phosphate limitation and so the eukaryotic phospholipid pathways could be maintained. This would indicate that the yellow isolates (RK1.3 and CM1.3Y) are *G. sulphuraria,* as there is a high quantity of phospholipids PC and PE but no betaine lipids are observed on either the 1D-TLC or the 2D-TLC. Although betaine lipids are more easily detected using LC/MS, which at the time of writing, has not been completed. For the green isolates (CM1.3G and KP1.2G), the high quantity of MGDG and DGDG than phospholipids (figs 3.13 and 3.14) and in comparison to the yellow isolates (fig 3.12) is evidence that these isolates are from the Cyanidiaceae family. As stated previously, the lack of reliable data for *Cyanidium* means it is difficult to use the polar lipid profiles to determine the taxonomy of the green isolates past the family level. Likewise there is a lack of data for *Galdieria* species, so it is unclear whether the lack of betaine lipids is significant or indicative of differences between *G. sulphuraria* and *G.phlegrea*. Although, the polar lipid profile in combination with the FAME profile is clear evidence that the yellow isolates are *Galderia* strains and the green isolates are from the family Cyanidiaceae. Notably there are very few observable differences between the two green isolates and between the yellow isolates. This is evidence that KP1.2G and CM1.3G are closely related. Likewise, CM1.3Y and RK1.3 are closely related despite being geographically distant.

3.4.5 Identification of the isolates

The combination of morphological, ecophysiological and biochemical data (table 3.6) is strong evidence that the yellow isolates are *Galdieria*, most likely *G. sulphuraria,* and that the green isolates are from the family Cyanidiophyceae.

The yellow isolates, RK1.3 and CM1.3Y, have many of the key diagnostic characteristics of *Galdieria,* namely facultative heterotroph, presence of α-linolenic acid and only being able to utilise ammonium as the sole inorganic nitrogen source. So are *Galdieria* and are likely *G. sulphuraria* as the NZ type IV strains isolated previously In New Zealand are *G. sulphuraria* based on the *rbc*L phylogeny [26] and *G.phlegrea* has only been observed in Italy and Turkey [77,159]. It would be difficult to determine if these strains are *G.phlegrea* with physical data, as no diagnostic characteristic have been determined [38,75,77] and it is difficult to distinguish between *Galdieria* without accurate phylogenetic data [36,77]

Evidence indicates the green isolates, CM1.3G and KP1.2G, are from the Cyanidiaphyceae family as no heterotrophic growth was observed and these isolates lack α-linolenic acid, however it remains difficult to determine the genus of these isolates. Initially, the oblong club-shaped juvenile cell morphology (figs. 3.6 and 3.7) identified that these isolates were from the *Cyanidioschyzon* lineage (fig. 2.1). However, later observations indicated the presence of endospores in both KP1.2G and CM1.3G (figs. 3.6; 3.7). This would potentially identify these isolates as abnormally shaped *Cyanidium*. There is evidence of *C. caldarium* in NZ based on a plastid 16S rRNA gene similarity [27] and the abnormal shape could be due to mutation relating to the cell wall.

However, Topin *et al.* 2008 describe the NZ type V strains and Japanese strains as "somewhat oblong" (fig. 2.5, H), and the phylogenetic analysis of the *rbc*L and 18S rRNA genes indicate that those strains were closely related to *G. maxima* IPPAS P507[26]. If the KP1.2G and CM1.3G are related to *G. maxima* it would be significant, as the type strain, *G. maxima* IPPAS P507, is a facultative heterotroph [36,61]. This led to the misidentification of *G. maxima* IPPAS P507 as a *Galdieria,* however later molecular and phylogenetic data would indicate that it was most closely related to the autotrophic *Cyanidioschyzon* lineage [28,29,38]. If related to *G. maxima*, isolates KP1.2G and CM1.3G as autotroph endospore producing Cyanidiales, could represent the link between *G. maxima* and *C. merolae*. This would require further molecular and phylogenetic data to be determined accurately.

While there are significant differences between the green and yellow isolates, there seems to be little variation in observed phenotype within the groups. All the yellow isolates have the shared observed cell morphology, as do all the green isolates. Even with the isolates that have been further characterised, there are only minor differences observed such as the difference in optimal pH between KP1.2G and CM1.3G (table 3.6). It is possible that the methodology lacks enough resolution to distinguish between isolates, however it is evidence that all the yellow isolates are *G. sulphuraria* and are closely related. Similarly, all the green isolates are closely related. This stresses the importance of the molecular and phylogenetic analysis in Chapter 4, without which accurate taxonomy could not be determined.

3.5 Conclusion

Phenotypic data gathered in this chapter supports the hypotheses that New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* and these lineage differentiated using phenotypic characterisation. However it lacks the resolution to determine taxonomy to a genera level therefore molecular data and phylogenetic analysis is needed

Chapter 4: Phylogeny Of NZ Cyanidiales

4.1 Introduction

4.1.1 Background

Molecular and phylogenetic data have been key in resolving the taxonomy of Cyanidiales. Cyanidiales phylogeny is often determined using the *rbc*L (Rubisco large subunit) sequence [26,32,38,68,78,160] and/or the 18S rRNA gene sequence [26,29,72]. The use of single protein-coding genes, particularly *rbc*L, can be prone to misidentification due to mutational saturation and variation in the third codon position [36], if the dataset is too small or the model does not account for it. There can also be conflicts between the phylogeny of nuclear and plastid genes [26]. Multigene concatenated datasets have been used to resolve the taxonomy of Cyanidiales [38] and Rhodophyta [28,29] to greater detail. A three plastid protein gene dataset consisting of *psa*A (photosystem I P700 chlorophyll A apoprotein A1), *psb*A (photosystem II reaction center protein D1) and *rbc*L, have been previously used to determine the Cyanidiales lineage [38]. There are two family lineages within the Cyanidiales; the Cyanidiaceae and the Galdieriaceae. *Galdieria sulphuraria* (*Galdieria* A) and *G.phlegrea* (*Galdieria* B) represent the two primary genera within Galdieriaceae family. The *Cyanidium caldarium* lineage, a mesophilic *Cyanidium* genus, the *Cyanidioschyzon* and *Galdieria maxima* lineages represent the genera within the Cyanidiaceae family (fig 2.3).

Molecular data for the New Zealand Cyanidiales is limited to only two studies in the 2000s [26,27]. Donachie *et al.*, (2002) described the plastid 16S rRNA gene sequence (99.1% similarity to *C. caldarium*) from a single strain isolated from White Island.

A more extensive sampling of Cyanidiales from NZ geothermal sites in Toplin *et al.* 2008 included sequencing of the 18S rRNA and *rbc*L genes. These isolates were separated into three groups using morphology and molecular data: NZ type IV, V and VI strains [26]. Based on *rbc*L phylogeny, NZ type IV strains were from the *G. sulphuraria* lineage and NZ type V and VI strains were from the *G. maxima* lineage. However, the 18S rRNA phylogeny places NZ type IV strains within the *G. maxima* lineage. This conflict between the *rbc*L and 18S rRNA phylogeny is unexplained at this point and is indicative of an issue with reliance on single-gene phylogeny. None of those New Zealand strains in Toplin *et al.* were identified as *C. caldarium*[26]. It is possible that the plastid 16S rRNA gene sequence misidentified the strain due to the conserved sequence and share similarity within a family. For example, it could be possible the plastid 16S rRNA gene from *G. maxima* or *Cyanidioschyzon merolae* would have similarity with Cyanidium strain.

Ten Cyanidiales representatives have been isolated from New Zealand geothermal sites as part of this study (Chapter 3, section 3.3.1). The isolates can be separated into yellow isolate and green isolates based the pigment of colonies when growth with exogenous sugars present. There are four “green” isolates and the six “yellow” isolates. Further characterisation of two green isolates (CM1.3G and KP1.2G) and two yellow isolates (CM1.3Y and RK1.3), determined that the green isolates are from the Cyanidiaceae family with the *G. maxima*/Cyanidioschyzon lineage and the yellow isolates are Galdieria species. For the phenotypic characterisation However, the phenotypic characterisation could not accurate determine the taxonomy of these isolates to the genera level hence phylogenetic analysis is required.

4.1.2 Aim of section

Aim of this Chapter is to generate molecular data to determine the phylogeny and classification of the ten isolates described in the previous chapter.

These data will (a) provide phylogenetic data to support chemotaxonimc characterisations (Ch x), determine if isolates represent novel strains within Cyanidiales, and (c) assist in clarifying the conflicting taxonomy presented in the Toplin *et al.* 2008 manuscript.

Three hypotheses will be tested (Section 2.6):

A) New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* lineage

B) The two New Zealand “*Galdieria*” lineages (*G. sulphuraria* and *G. maxima*) can be differentiated via phenotypic characteristics and phylogenetic analysis.

C)The 16S sequences of my isolates will have a close similarity (>99% pairwise identity) to *C. caldarium*

To address these hypotheses, I will sequence a set of five genes from each of the ten isolates. The genes sequenced will be; the nuclear 18S rRNA gene, the plastid 16S rRNA (* see footnote in next paragraph), *psa*A, *psb*A and *rbc*L genes. The BLAST search for the plastid 16S rRNA sequence will determine if the isolates have high similarity to *C. caldarium* or the previously isolated Cyanidiales NZ4 strain from White Island [27]. Single gene phylogenetic trees of the 18S rRNA and *rbc*L genes will be calculated to determine relation to NZ strains previously isolated in Toplin *et al.* 2008. Phylogeny will also be calculated from two concatenated dataset; a three plastid protein gene set (*psa*A, *psb*A and *rbc*L) as used in Ciniglia *et al.* 2004 and five gene set of all genes sequenced (18SrRNA, 16S rRNA, *psa*A, *psb*A and *rbc*L) as used in Yoon *et al.* 2006.

* (Footnote). I originally intended that the five gene data set would include sequence from the nuclear calmodulin (*CaM*) gene rather than the plastid 16SrRNA gene. The *CaM* gene sequence has been used to determine phylogeny of Cyanidiales previously [75,79,161], so published sequences would be available from the GenBank. This would mean that the dataset would consist of two nuclear genes and three plastid genes, which could be able to indicate if the plastid phylogeny was different from the nuclear phylogeny. Despite troubleshooting efforts (see section x.x for details), the *CaM* primers did not produce a successful PCR reaction, as such it was decided to use the plastid 16S rRNA sequence instead, as the sequences had been generated and a similar five-gene dataset had been used before to determine the phylogeny of Rhodophyta [29].

4.2 Methods

4.2.1 DNA extraction

4.2.1.1 Biomass generation

All ten isolates were grown mixtrophically on MAll-Ht plates,Modified Allen's medium [136] pH 2.5 using 1.5 % (w/v) phytagel,supplemented with 3gL^{-1} glucose and 3gL^{-1} glycerol (see section x.x for details) for generation of biomass for DNA extraction. Cultures were incubated at 45°C under constant illumination with a warm white fluorescent lamp. Incubation for 2-3 weeks generated sufficient biomass for DNA extractions.

Two - four loopfuls of cells were aseptically scraped from the MAll-Ht plates using 5µL disposable inoculating loops and then resuspended in 1mL of sterile MillQ water. Cells were then centrifuged (11,000 xg for 1 min), the supernatant removed, and resuspended again in 150µL of sterile MillQ water. This cell suspension was used in the extraction step.

4.2.1.2 DNA extraction

Chloroplast and genomic DNA were extracted and isolated using the NucleoSpin® Soil kit (Macherey-Nagel, Düren, Germany). This kit has the choice of two lysis buffers (SL1 and SL2) which can be used with or without the Enhancer solution (SX), in preliminary tests of the four possible extraction condition combinations, the SL2+FX gave the most constant yield and worked for all 10 isolates. Mechanical lysis was achieved with a Bead-mill at 5 ms^{-1} for 30 seconds. Final elution protocol was done in two subsequent elution steps with 25 μL fresh elution buffer each step for final extraction volume of 50 μl, this is an alternative method for increasing DNA concentration suggested in the Manual. UV-VIS spectroscopy was used to determine the concentration of extracted DNA (absorbance at 260 nm) and its purity (absorbance ratio 260/230 and 260/280) using NanoDrop (Thermo Fisher Scientific, Waltham, Massachusetts, EUA), The above protocol would, with sufficient biomass, would result in a 50μL yield with minimum DNA concentration of 40ng/μl.

4.2.2 Amplification of gene targets via PCR

Table 4.1 Primers used in final PCR reactions

Target Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Reference
18S rRNA	**CdmF** GTCAGAGGTGAAATTCTTGGATTTA	**CdmR** AAGGGCAGGGACGTAATCAACG	Gross *et al.* 2001
***rbc*L**	**RbcL1F** AACCTTTCATGCGTTGGAGAGA	**RbcL1R** CCTGCATGAATACCACCAGAAGC	Toplin *et al.* 2008
***psa*A**	**psaA130F** AACWACWACTTGGATTTGGAA	**psaA1760R** CCTCTWCCWGGWCCATCRCAWGG	Yoon *et al.* 2002
***psb*A**	**psbA-F** ATGACTGCTACTTTAGAAAGACG	**psbA-R1** GCTAAATCTARWGGGAAGTTGTG	Yoon *et al.* 2002
16S rRNA	**9F** AGAGTTTGATCMTGGCTCAG	**1492R** GGHTACCTTGTTACGACTT	

Target genes from each of the 10 Cyanidiales isolates were amplified via PCR. A comprehensive list of primers targeting the *rbc*L *psa*A, *psb*A, *CaM*,18S rRNA and plastid 16S rRNA genes listed in table 6.4. A list of the primers that chosen after initial testing are listed in table 4.1.

PCR reactions contained the following components: 10μl of MyFi Mix 2x (Bioline,U.S.A.), 2μl forward primer (10μM), 2μl reverse primer (10μM), ≥200ng of Template DNA (5μL at 40ng/μl) and made up to 25μL with PCR Grade water.

Primer choice and Cycling conditions were optimised for target sequence and the Cyanidiales. final cycling conditions are listed in table 4.2. For most target sequences, such as *rbc*L, 16S rRNA and 18S rRNA, cycling conditions were determined by following the specification of the

MyFi master mix, then sourcing the optimal annealing temperature from the literature [26, 27, 72] and adjusting the extension time to be appropriate for the length of the target sequence as per specification of the MyFi mix (longer sequences require longer extension times).
For some sequences no anneal temperature is given in the literature and there were multiple primer sets, as these primers were designed application in broad range of species. In these cases, first the 2 primers set that would give longest length were trialled in PCR reactions with an anneal temperature of 55°C with appropriate extension time using two extracts one from yellow isolate and one green isolate. The primer set that indicate positive results (gel electrophoresis) with both extracts was used. If no results were detected the primer set that would give the longest sequence were trialled in a gradient PCR with annealing temperatures from 45-60°C at 5°C increments.

Table 4.2 Final Cycling conditions used for PCR reactions

Genes	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Cycles
18S rRNA and *rbc*L	95°C for 3 min	95°C for 15s	55°C for 15s	72°C for 25s	72°C for 2 min	30
*psb*A	95°C for 3 min	95°C for 15s	55°C for 15s	72°C for 45s	72°C for 2 min	30
*psa*A and 16SrRNA	95°C for 3 min	95°C for 15s	50°C for 15s	72°C for 45s	72°C for 2 min	35

Integrity of the PCR products was confirmed by gel electrophoresis with 1% (w/v) agarose gel. Clean-up of PCR products was done using a DNA Clean & Concentrator Kit (Zymo Research, U.S.A.). Cleaned PCR products were sequenced using Macrogen Ez-Seq (Seoul, South Korea), in both forward and reverse directions.

4.2.3 Assembly of Sanger sequencing reads

With the raw sequence data, the forward and reverse reads were separate hence a consensus sequence need to be generated. Sequence data was imported, managed and edited using Geneious Prime 2019.1.3 (www.geneious.com). To remove poor quality base calls from the reads both the 5’ and 3’ ends were trimmed, using the Geneious trim tool with the Error probability limit set to 0.5%. The forwards and reads of a sequence were then assembled using the Geneious assembler set to highest sensitivity. The assemblies were checked for disagreements and were manually where appropriate. The consensus sequences were extracted from the assemblies based on highest call quality. These consensus sequences were used in the subsequent analyses.

4.2.4 Similarity of the plastid 16S rRNA gene sequences.

To determine if any of the 16S rRNA sequences from the ten isolates shared high similarity with a Megablast search [167,168] (excluding results from uncultured/environment samples) was run within Geneious. The BLAST hit table was sorted by highest "Grade," which is a percentage calculation combining the query coverage, e-value and identity values by Geneious [169].

4.2.5 Phylogenetic analysis of single-gene dataset 18S rRNA and *rbc*L

Aim was to replicate methods in Toplins *et al.* (2008) to allow comparison to sequences from previously isolated NZ Cyanidiales strains. Done within the Geneious ecosystem unless otherwise stated.

4.2.5.1 Phylogenetic analysis of *rbc*L gene

The 10 partial *rbc*L sequences from the samples were aligned using ClustalW2.1 [162] , penalties were 15 for gap open, 6.66 for gap extension and free end gaps. The correct reading frames of the sequences was determined by running a BLASTX[167] search. The sequences in alignment were manually edited to start on with a 1st codon positioned nucleotide and end with a 3rd codon position nucleotide. The nucleotide sequences were then translated using Geneious translate tool with the bacterial translation table. From the 10 samples this generated a set of 10 nucleotide sequences and set of 10 protein sequence.

Published sequence data from Toplin *et al.* 2008 and other Rhodophyta phylogeny[28, 29, 38] were obtained from GenBank (www.ncbi.nlm.nih. gov). Accession numbers in table 6.5 (section 6.2). The DNA sequences were aligned ClustalW2.1 [162] and the protein sequences ClustalW2.1 using the BLOSUM cost matrix. This generated a protein dataset and DNA dataset consisting of 87 sequences

The phylogeny of the DNA dataset was calculated using Bayesian inference which was done with Mrbayes 3.2.6 [165,166] using the GTR+G evolutionary model with *Bangia fuscopurpurea*

SAG59.81(AY119771.1) as the outgroup. Analyses were run for 1,000,000 cycles with subsampling every 100 cycles and a burn-in length of 50,000 cycles. To avoid overparameterization, priors probability distributions were left as default. GTR+G was determined as suitable evolutionary models using Jmodeltest 2.1.10 [163,164] with AIC outside Geneious ecosystem.

The phylogeny of the Protein dataset was calculated using Maximum likelihood methods with was done with PhyML 3.3.2 [164] with the JTT substitution model and 200 bootstrap replicates. Model was used in previous ML analysis of Cyanidiales *rbc*L protein sequences[28, 38].

4.2.5.2 Phylogenetic analysis of 18S rRNA gene

Published sequence data from Toplin *et al.* 2008 and other Rhodophyta phylogeny papers [29, 72] were obtained from GenBank (www.ncbi.nlm.nih. gov). Accession numbers in table 6.6 (section 6.2). The DNA sequences from the literature were aligned with the 10 DNA sequences from the samples using the Geneious alignment tool with the 70% similarity (IUB) cost matrix and set to automatically determine optimal directions for alignment. Once the correct read directions were set, the dataset was realigned using Clustalw2.1 and manually trimmed to a maximum length of 800bp. The final dataset consisted of 62 sequences. The phylogeny of the DNA dataset was calculated using Bayesian inference which was done with Mrbayes 3.2.6 [165,166] using the HKY+I+G evolutionary model, using *Rhodochaete parvula* (AF139462.1) as the outgroup. Analyses were run for 1,000,000 cycles with subsampling every 100 cycles and a burn-in length of 50,000 cycles. To avoid overparameterization, prior probability distribution were left as default. HKY+I+G was determined as suitable evolutionary models using Jmodeltest 2.1.10 [163,164] with AIC outside Geneious ecosystem.

4.2.6 Phylogenetic analysis of the concatenated datasets

Aim was to replicate the three-plastid protein dataset from Ciniglia *et al.* 2004 that was used to define Cyanidiales lineages[38] and the Five-gene dataset from Yoon *et al.* 2006 that was used to define lineages within Rhodophyta

4.2.6.1 Three plastid protein concatenated dataset

Published sequences[28,29,38] were obtained from GenBank (www.ncbi.nlm.nih. gov). Accession numbers in table 6.7 (section6.2). For the concatenated datasets, the individual gene dataset had to be aligned and trimmed to a consistent length separately. The individual gene datasets were aligned separately using ClustalW2.1 with the BLOSUM cost matrix and trimmed to the to the same length manually. The DNA sequences were concatenated using the Geneious software, realigned using ClustalW2.1 with the BLOSUM cost matrix and a translated dataset was generated using Geneious translate tool with the bacterial translation table.

The complete the three plastid dataset, the dataset consisted of 41 sequences 2250bp or 750aa in length, from *rbc*L (456bp; 97aa), *psa*A (921bp; 307aa) and *psb*A (873bp; 291aa) datasets

The phylogeny of the DNA sequences was calculated using Bayesian inference which was done with Mrbayes 3.2.6 [165,166] using the GTR+I+G evolutionary model with *Bangia Subsimplex* as the outgroup. Analyses were run for 1,000,000 cycles with subsampling every 100 cycles and a burn-in length of 50,000 cycles. To avoid overparameterization, priors probability distributions were left as default. GTR+I+G was determined as suitable evolutionary models using Jmodeltest 2.1.10 [163,164] with AIC outside Geneious ecosystem.

The phylogeny of the Protein sequences was calculated using Maximum likelihood methods with was done with PhyML 3.3.2 [164] with the JTT substitution model and 1000 bootstrap replicates. Model was used in previous ML analysis three-plastid protein dataset from Ciniglia et al. 2004. The third codon positions of the protein-coding sequences are prone to mutational saturation [28,36]. To determine if this affected the topologies of the Bayesian inferred trees, using the Geneious masking, the third codon positions were removed from the DNA sequence alignment. The phylogeny of this edited datasets was calculated using the same Bayesian inference where the model and outgroup choice was not changed.

4.2.6.2 Five gene concatenated dataset

Published sequences[28,29,38] were obtained from GenBank (www.ncbi.nlm.nih. gov). Accession numbers in table 6.8 (section 6.2). For the concatenated datasets, the individual gene dataset had to be aligned and trimmed to a consistent length separately. The individual gene datasets were aligned separately using ClustalW2.1 and trimmed to the to the same length manually. The DNA sequences were concatenated using the Geneious software, realigned with ClustalW2.1.

The complete 5 gene dataset consisted of 21 sequences 4399bp in length, from 18S rRNA(474bp), *rbc*L (456bp), *psa*A (1242bp) *psb*A(918bp) and plastid 16SrRNA (1309bp) datasets.

The phylogeny of the DNA sequences was calculated using Bayesian inference which was done with Mrbayes 3.2.6 [165,166] using the GTR+G evolutionary model with *Bangia Subsimplex* as the outgroup. Analyses were run for 1,000,000 cycles with subsampling every 100 cycles and a burn-in length of 50,000 cycles. To avoid overparameterization, priors probability distributions were left as default. GTR+G was determined as suitable evolutionary models using Jmodeltest 2.1.10 [163,164] with AIC outside Geneious ecosystem.

The phylogeny of the DNA sequences was also calculated using Maximum likelihood methods with was done with PhyML 3.3.2 [164] with the GTR substitution model and 1000 bootstrap replicates. The third codon positions of the protein-coding sequences are prone to mutational saturation [28,36]. To determine if this affected the topologies of the Bayesian inferred trees, using the Geneious masking tool, the third codon positions were removed from the *rbc*L, *psa*A and *psb*A sequences in the alignment. The phylogeny of this edited datasets was calculated using the same Bayesian inference where the model and outgroup choice was not changed.

4.2.7 Genome Extraction

Genomic extractions of KP1.2G and RK1.3 cyanidiales isolates were undertaken using a modified CTAB method as described previously Healey *et al*.[170]. Briefly, this extraction was conducted in the following manner. Cell lysing was achieved by grinding the frozen biomass (with liquid nitrogen) in a mortar and pestle. Cell lysation was suspend in 10mL of the extraction buffer (100 mM Tris-HCl (pH 7.5), 25 mM EDTA, 1.5 M NaCl, 2% (w/v) CTAB, and 0.3% (v/v) β-mercaptoethanol- added immediately before use) in a 50mL centrifuge tube and incubated at 65°C for 1 hour. Cell debris removed by centrifugation (5 min at 5000 × g) and pouring supernatant into a new 50mL centrifuge tube. The Protein extraction step was undertaken adding 10mL of chloroform:isoamyl alcohol (24:1) and mixing by inversion for 5 mins. The samples were centrifuged (10 min at 5000 × g) to separate layers and organic phase was discarded. Then 10μl of RNAse A (10mg/mL) was added to the aqueous phase which is incubated at 37°C for 15 min. This was followed by another protein extraction step. The aqueous from that protein extraction was added with 5mL of 5 M NaCl solution to 30mL of pre-chilled 95% ethanol. This solution was incubated for at -20°C for 1 hour. The DNA was then pelleted (centrifugation for 5 min at 5000 × g) and washed with 70% ethanol. Centrifuged again (5 min at 5000 × g), supernatant was discarded, and the pellet is allowed to air-dry. Finally, the DNA was suspended in 200μL of TE buffer.

The Cyanidiales isolates KP1.2G and RK1.3 were chosen as representatives of the green and yellow isolates respectively. Biomass from both strains were grown in 1L Schott bottles with 700mL 2xV4 medium[23,135]. The headspace of the bottle enriched with CO_2 by purging with a continuous stream of filter sterilised pure CO_2. Medium for the RK1.3 was supplemented with $5gL^{-1}$ of glucose to increase growth rate. All were grown at 45°C under constant illumination. Approximately 1.5g of biomass (wet weight) was harvested for each isolate. Biomass was aliquoted into three extractions with 0.5g for both strains, a total of six extractions. Gel electrophoresis and UV-vis spectroscopy (Nanodrop) methods were used to determine DNA yield, purity and integrity for each extraction.

4.3 Results

4.3.1 Genome extractions

The modified CTAB extraction method was partially successful as the extraction had acceptable DNA yields and very large fragment sizes, but the purity of the DNA was low. The bands of the KP1.2G are all above the final marker (10Kbp) with little shearing, indicating good integrity and fragment size above 10Kbp. However, the 260/230nm absorbance ratio was ~0.5 for all replicates, which indicates high levels of RNA contamination. The 260/280nm absorbance ratio was between 1.8 and 1.9 which indicates minimal protein contamination.

The bands for the RK1.3 extractions are also all above the final 10Kbp marker, however the a and c aliquots are streaky and the b band is well defined by the concentration of DNA is less. The aliquots for RK1.3 had absorbance ratios of ~2.1 for 26/280nm ratio and ~1.5 for the 260/230nm but had a clear peak at 260nm. Therefore the DNA yield was high but the purity was low.

It is possible that the RK1.3 aliquots were incubated at -20°C for too long during the precipitation stage, resulting in the precipitation of other solutes in the extraction such as NaCl. Given the cost and stringent DNA quality requirements for NGS, it was decided that genome extractions would need to be reattempted. The current method would need to be optimised, for example, using a greater concentration of RNase to remove RNA contamination or an alternative method would need to be used.

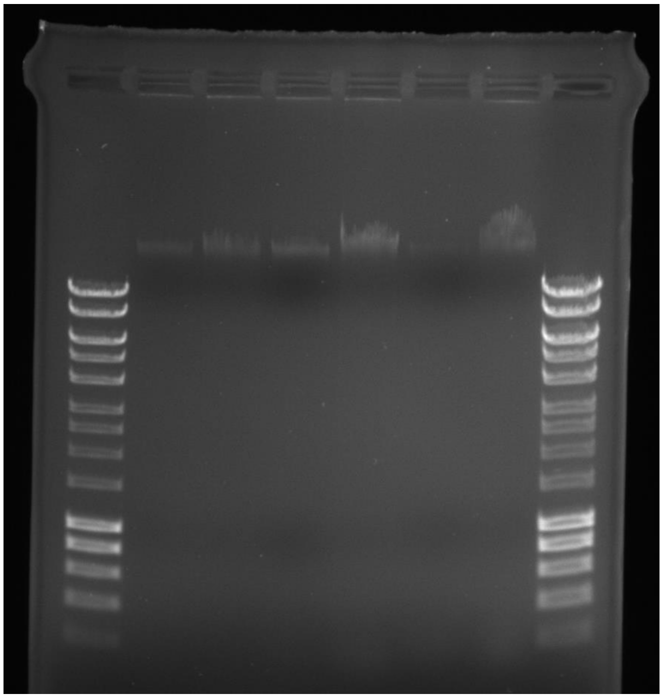


Figure 4.1: Gel electrophoresis result of the genome extractions. Lanes from left to right are; Hyperladder^{TM} 1kb (Bioline); KP1.2G extractions aliquots a, b, c; RK1.3 extractions aliquots a, b, c; Hypperladder^{TM} 1kb. Highest marker for the ladder is 10,037bp

4.3.2 The single-gene trees, *rbc*L and 18S rRNA

The *rbc*L Bayesian phylogenetic tree (fig4.2) has nearly identical topology to the Toplin *et al.* 2008 *rbc*L tree (fig 2.6), this will allow for classification of the isolates as the different type strains. The only unexpected topology is that the *Cyanidium* lineages form a sister clade with the “true” *Galdieria* clade rather than *Cyanidioschyzon/G. maxima* clade. However, that does not affect the topology of any of the isolates. All the green strains (C1G, C2G, C3G, KPG in fig.x.x) cluster closely within the NZ Type V clade (purple in fig 4.2). This clade forms a monophyletic group with the Japan Type III strains and *G. maxima* IPPAS P507, the type strain for *G. maxima* [61]. The single NZ type VI strain is also monophyletic with the *G. maxima* clade, although it is the most distant. All the yellow isolates (C2Y, C3Y, KPY, RK3 in fig.x.x), excluding RK2, cluster with the YNP type II clade which is sister to *G. sulphuraria* SAG108.79, an isolate from YNP [36,48].

This clade forms a monophyletic group with the NZ type IV clade separate from the Russian clade of *G.partita* and *G.daedala*, suggesting the yellow isolates and NZ type IV share an American lineage. The RK2 isolate separates from the other yellow isolates and forms a clade with the *G. sulphuraria* strains DBV021 and DBV015, which were isolated from Italian sites [38]. No isolates grouped with the NZ type VI strain (CCMEE 5713 WTP)

The *rbc*L ML phylogenetic tree (fig 4.3) is less resolved than the Bayesian tree. RK2 still forms a clade with *G. sulphuraria* strains DBV021 and DBV015. The green strains are also clustered within the NZ type V clade. The NZ type VI strain has formed a clade with the 5720WI strain within the NZ type V and Green isolate clade, however with extended branch length. This indicates that the evolutionary distance between the NZ type V and VI strains is great enough to justify distinct strains, but NZ type VI is more closely related to NZ type V strains than the isolates for Japan.

While the yellow isolates (excluding RK2) cluster together, the NZ type IV formed a sister clade with YNP type II and *G. Sulphuraria* SAG108.79, however the bootstrap proportions for the whole *Galdieria sulphuraria* clade are low.

For the 18S rRNA Bayesian tree (fig 4.4), the *G. sulphuraria* clades bifurcates into clear American and Eurasian lineages. The Eurasian clade includes the isolate RK2 which is sister to a clade of Russian strains, *G.partita* IPPAS P500 and *G.daedala* IPPAS P508 [61]. The American clade includes the other yellow strains, clustered together, and the YNP type II clade. The green isolates cluster within a large unresolved clade that includes; NZ type IV and V strains, Japan type III-A and B strains, and *G. maxima* IPPAS P507.

There is strong evidence that all of the green isolates are examples of NZ type V strains isolated previously in Toplin et al 2008. The isolate RK2 is not closely related to any strains isolated from NZ within this thesis or previously [26].

The yellow isolates are more closely related to the YNP type II and America strains than the NZ type IV strains, based on the bayesian inferred trees (Figs. 4.2, 4.4). However, with the ML inferred tree of the *rbc*L protein sequence (Fig 4.3), the NZ type IV strains are more closely related to American strains than yellow isolates were. The yellow isolates, unlike NZ type IV, do not collapse into the *G. maxima* clade in the 18S rRNA tree (fig 4.4).

Based on the *rbc*L, it can be concluded that NZ type IV and the Yellow isolates are *G. sulphuraria*, and that there is a common lineage.

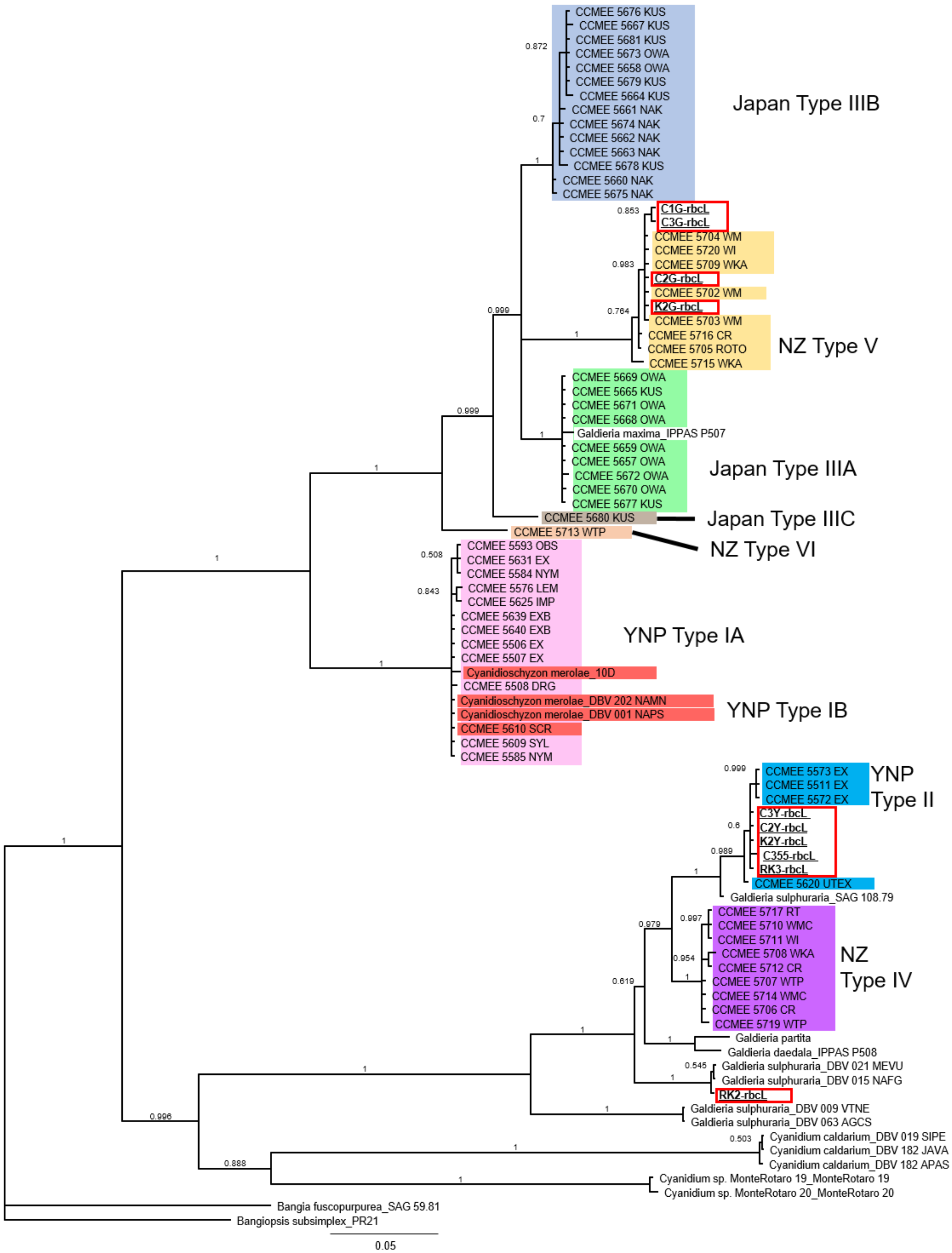


Figure 4.2: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the rbcL DNA sequence. Bayesian posterior probabilities are shown on above the branches as decimals. Sequences from this thesis are indicated with red boxes and sequences from Toplin *et al.* 2008 are coloured and labelled.

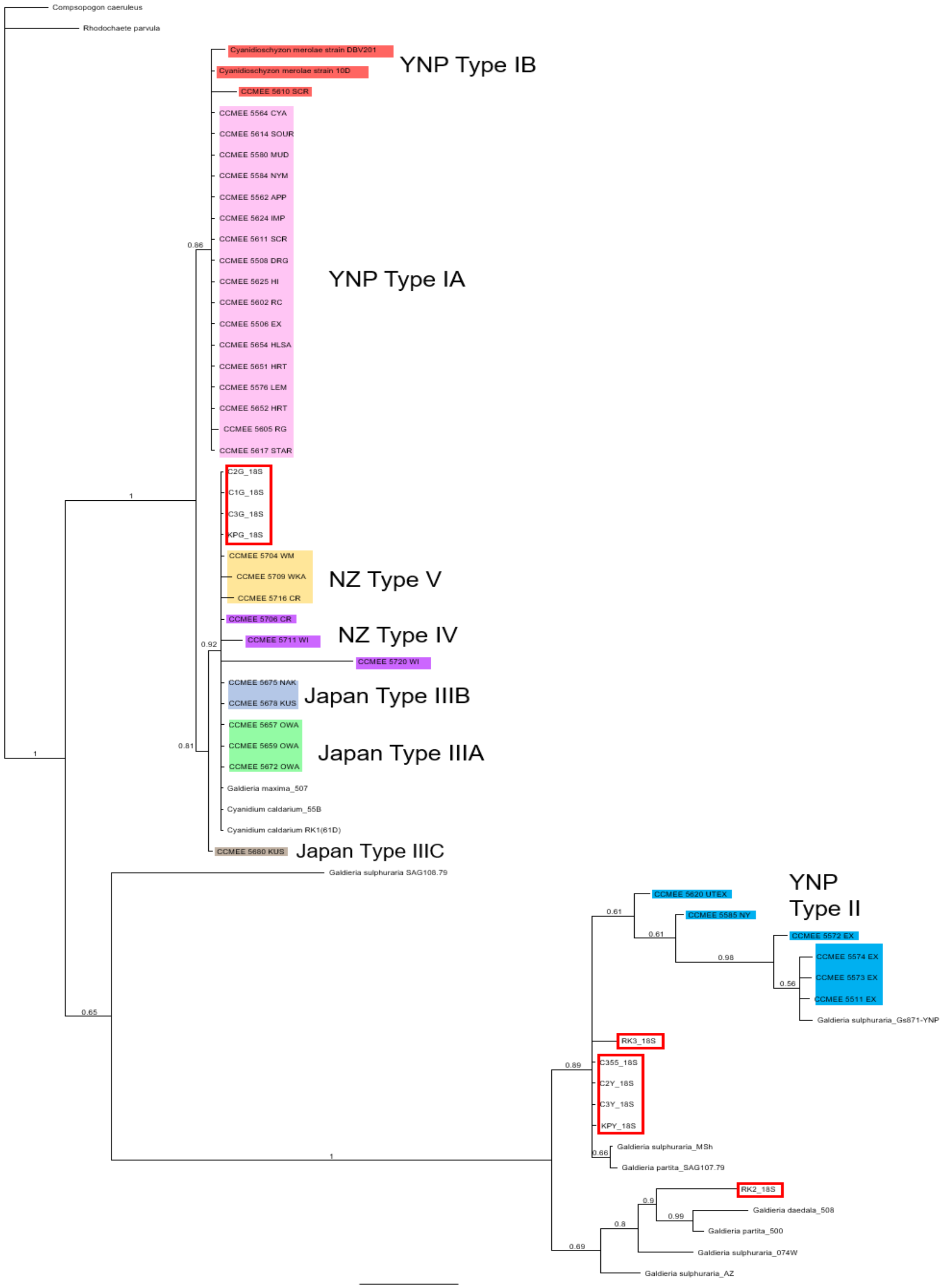


Figure 4.3: The phylogeny of New Zealand Cyanidiales inferred from maximum likelihood (ML) analysis of the rbcL protein sequence. Values from Bootstrap anaylsis are shown on above the branches Sequences from this thesis are indicated with red boxes and sequences from Toplin *et al.* 2008 are coloured and labelled.

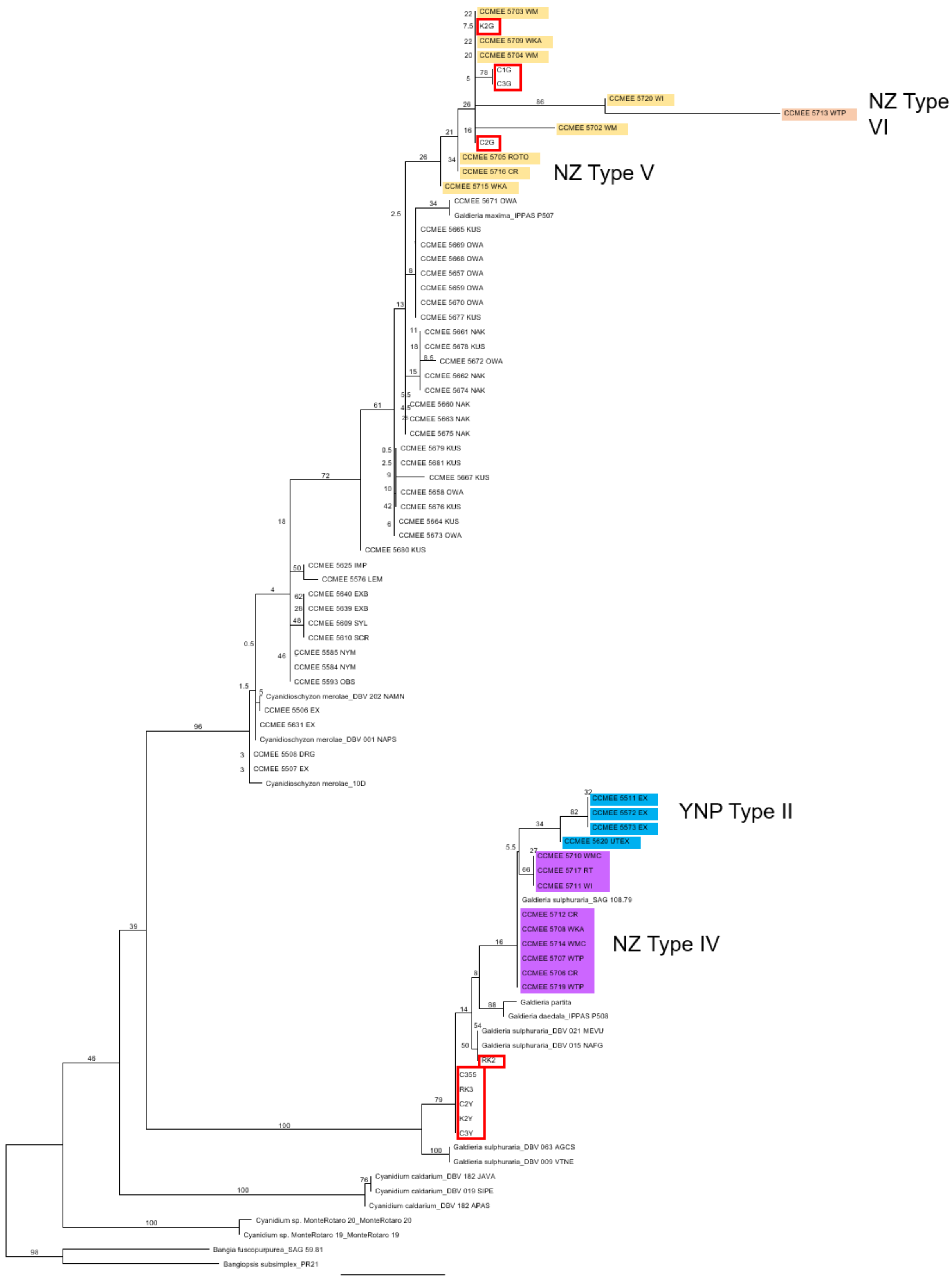


Figure 4.4: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the 18S rRNA gene sequence. Bayesian posterior probabilities are shown on above the branches as percentages. Sequences from this thesis are indicated with red boxes and sequences from Toplin *et al.* 2008 are coloured and labelled.

4.3.3 Concatenated Dataset and Placement with Cyanidiales

For the three-plastid gene dataset, there is consensus between the ML and Bayesian inferred phylogenetic trees (Figs 4.5, 4.6). All the green isolates cluster together within a clade, which is sister to the *G. maxima* clade and contains strains from Russia and Turkey [61,159].
Isolate RK2 forms a clade with the Italian *G. sulphuraria* strains [38].
All the other yellow isolates cluster together to form an unresolved clade with the America *G. sulphuraria* strain SAG108.79 [48].
The *Galdieria* B/*G.phlegrea* lineage forms a clade separate from the *Galdieria* A/*G. sulphuraria* lineage and does not contain any of the isolates from this thesis. The family Cyanidiaceae, has been resolved with separate *Cyanidium* lineages and *G. maxima/C. merolae* lineage. The only conflict between the trees is in the ML tree the RK2/Italian clade is moved to form a monophyletic clade with Eurasian *G. sulphuraria* clade; which given geographic distances, seems more probable. The Bayesian inferred tree from the dataset with the 3rd codon removed(fig 4.7), has identical topology to the ML protein sequence tree (fig 4.6, except for branch length). Indicating conflict is probably due to variability in the 3rd codon position.

For the Five-gene dataset, there is consensus between ML and Bayesian inferred trees (fig 4.8 and 4.9). The green isolates cluster together and form a clade with the *G. maxima* IPPAS P507. The yellow isolates, excluding RK2, all cluster together and RK2 separates from the other yellow isolates forming a clade with *G. sulphuraria* 074W, a strain from Java [91].
There are only minor changes in topology when the 3rd codon position is moved (Sfig 6.1).

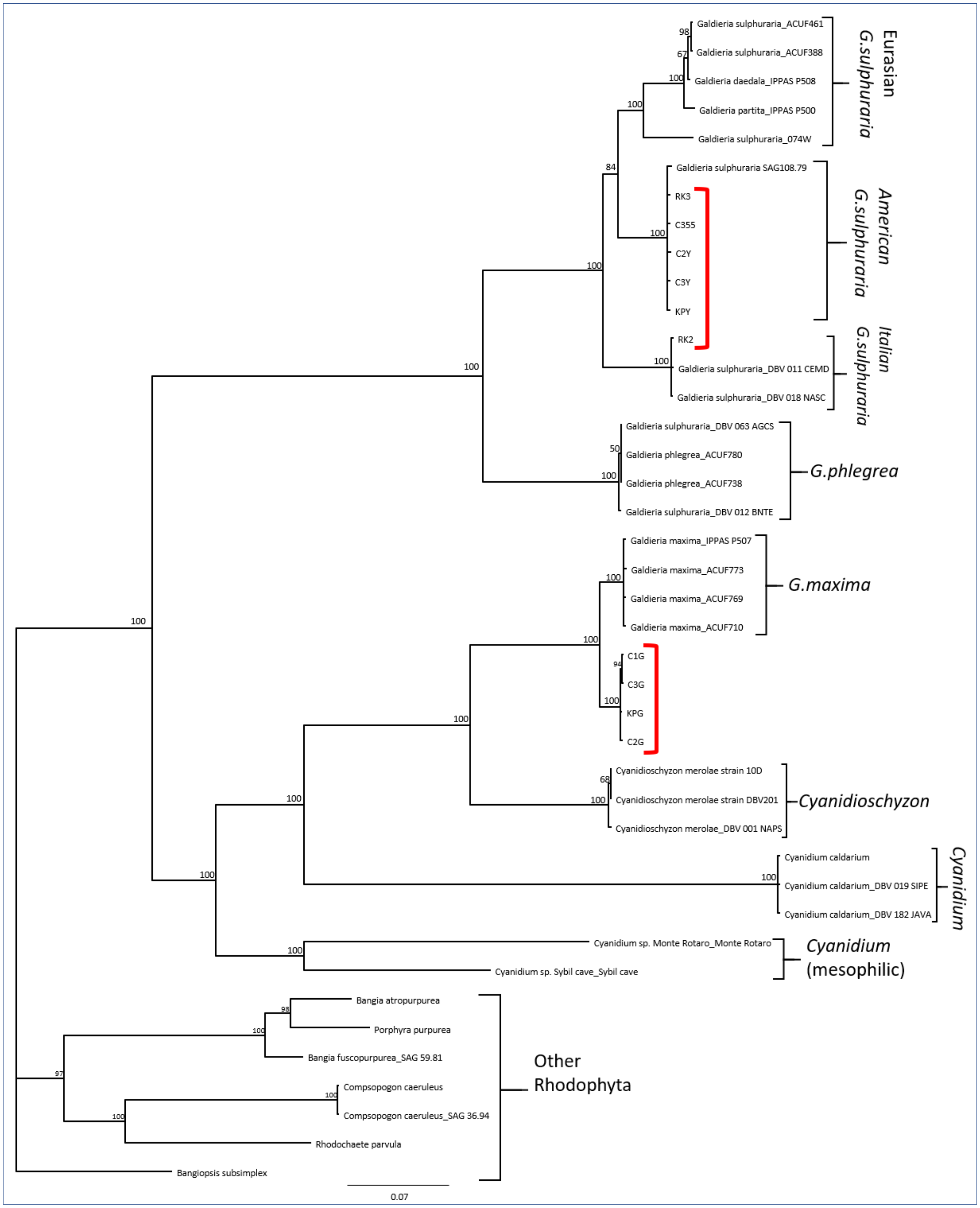


Figure 4.5: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the combined plastid DNA sequences of *psa*A, *psb*A and *rbc*L. Bayesian posterior probabilities are shown on above the branches as percentages. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets

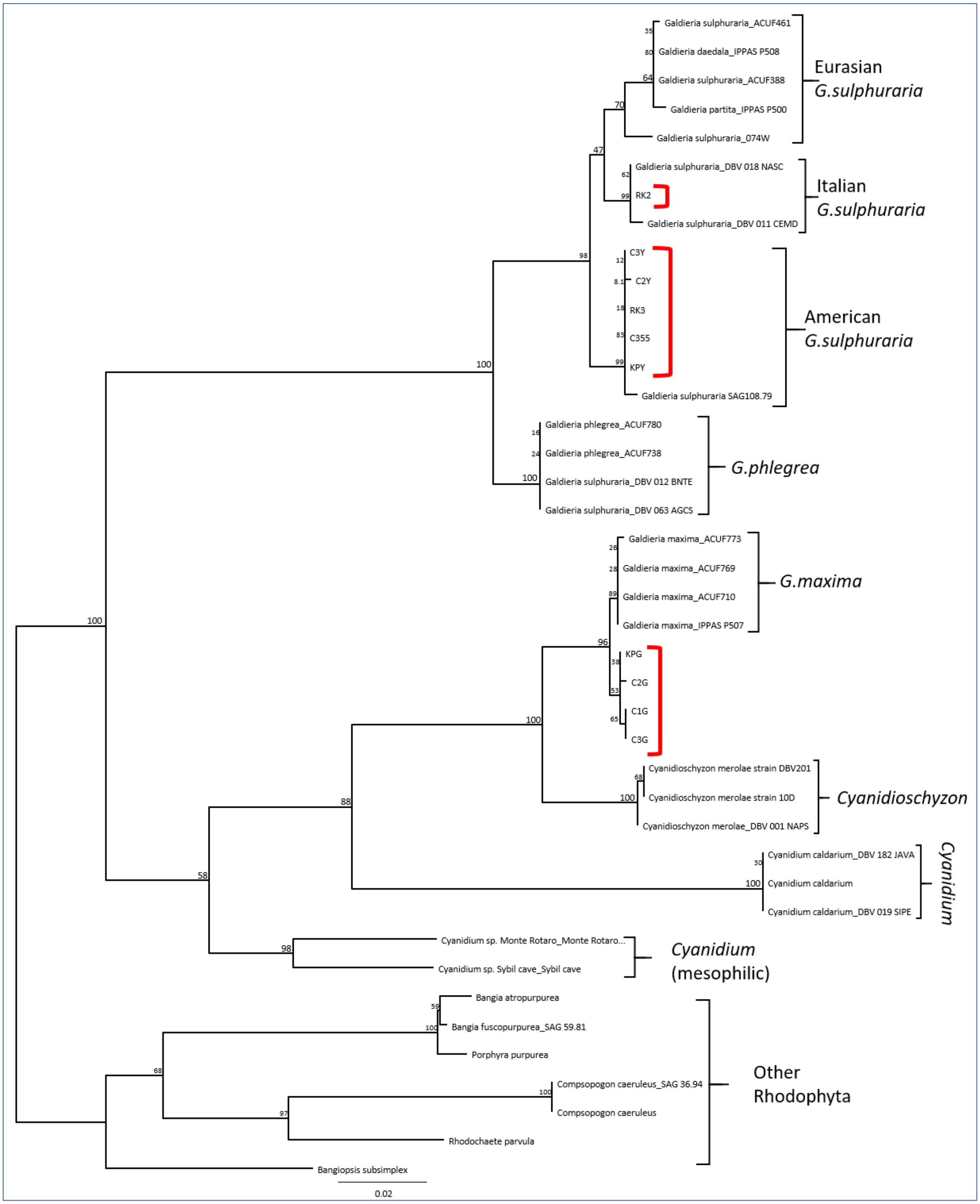


Figure 4.6: The phylogeny of New Zealand Cyanidiales inferred from maximum likelihood (ML) analysis of the combined plastid protein sequences of *psa*A, *psb*A and *rbc*L. Values from bootstrap analysis are shown on above the branches as percentages. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets

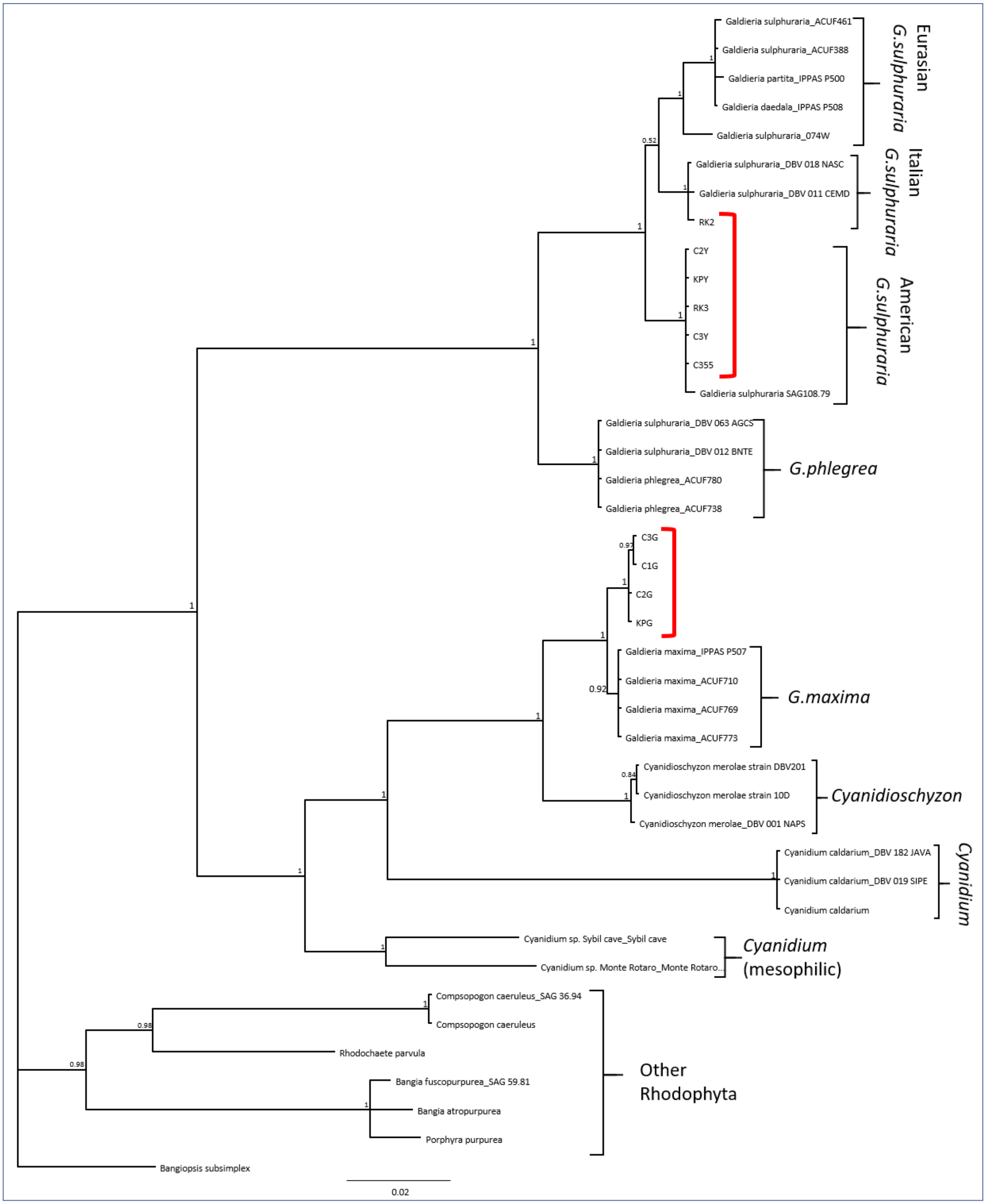


Figure 4.7: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the combined plastid DNA sequences of *psa*A, *psb*A and *rbc*L excluding third codon positions. Bayesian posterior probabilities are shown on above the branches as decimals. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets.

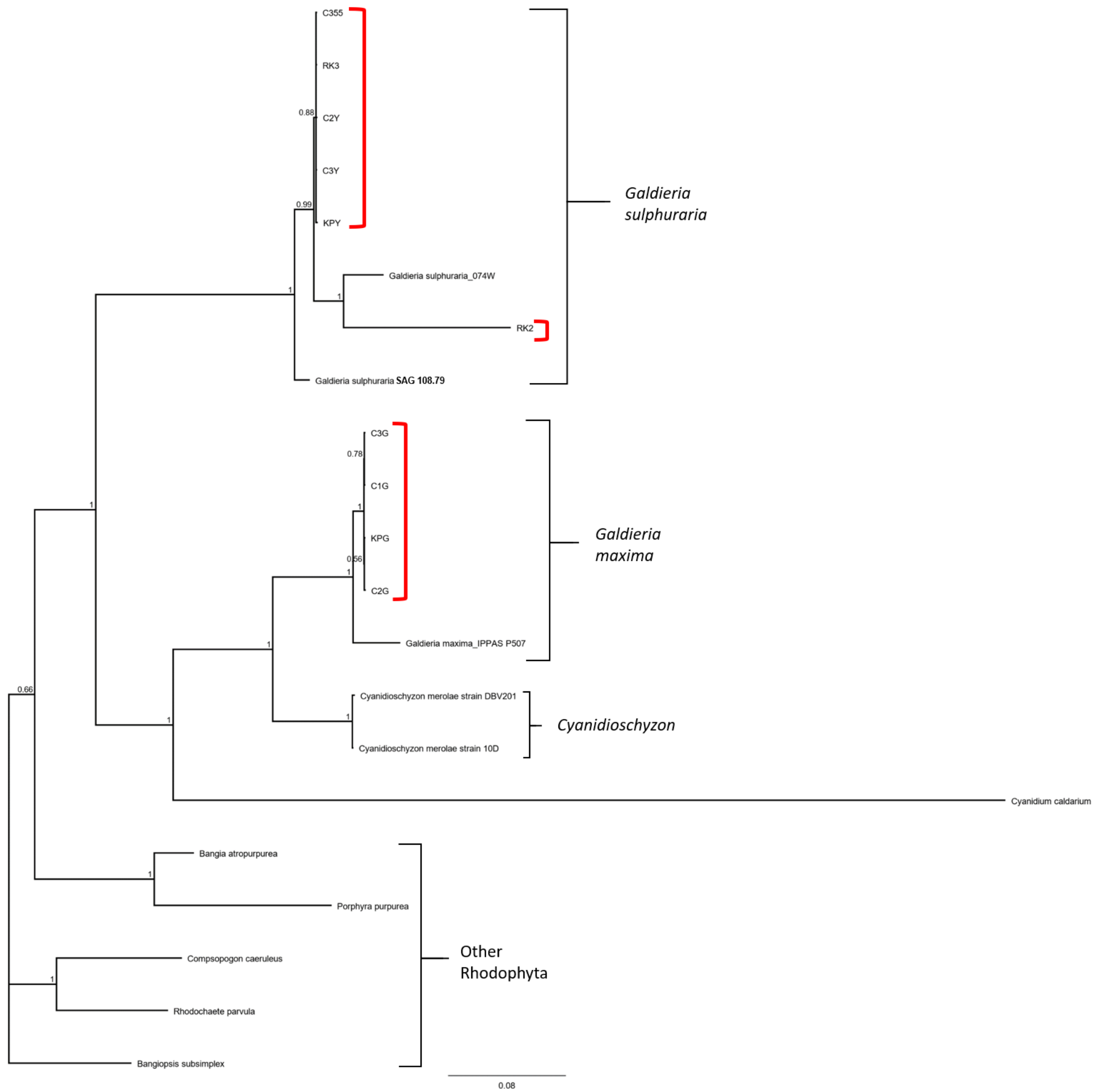


Figure 4.8: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the combined DNA sequences of *psa*A, *psb*A, *rbc*L, 18S rRNA and plastid 16SrRNA genes. Bayesian posterior probabilities are shown on above the branches as decimals. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets

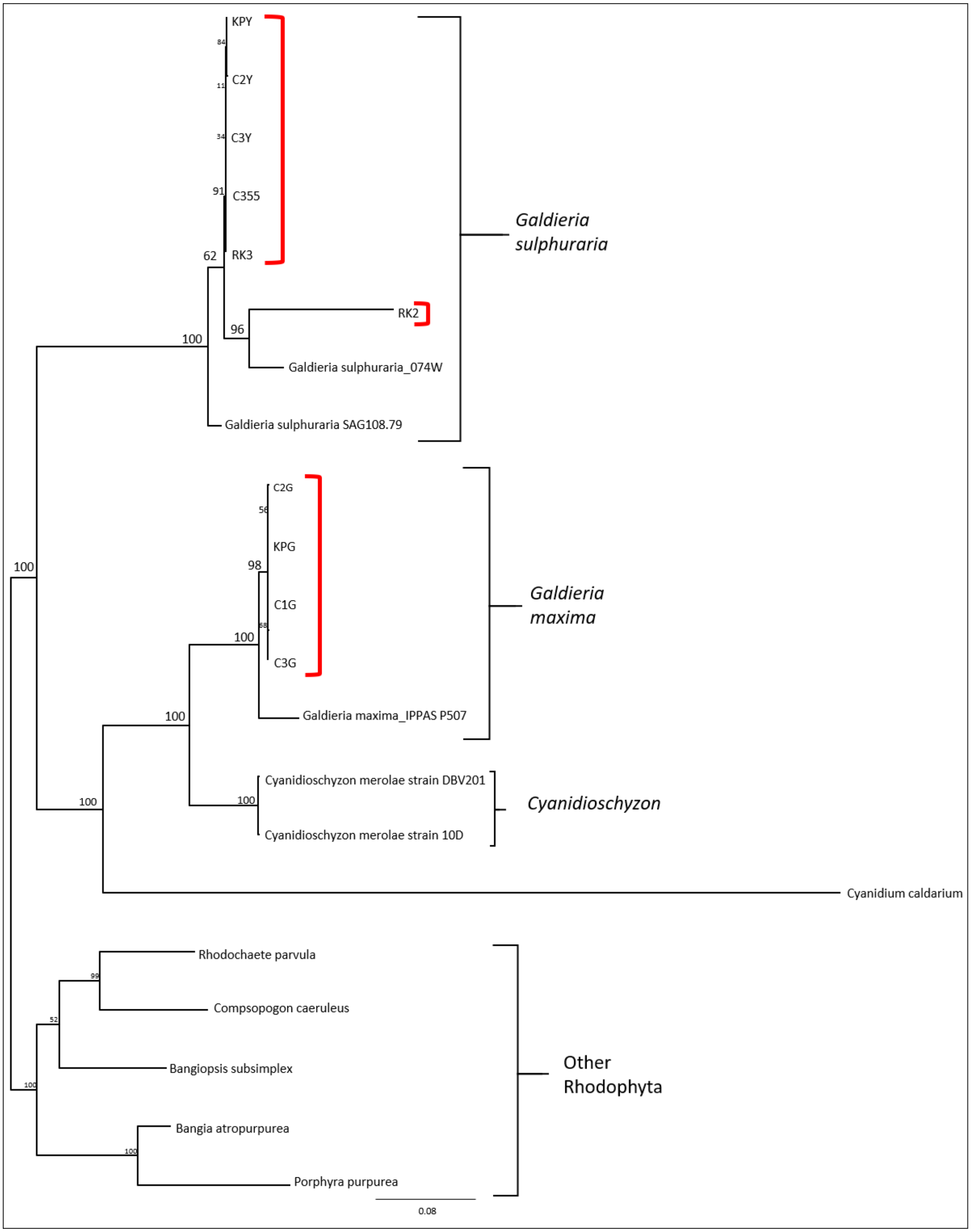


Figure 4.9: The phylogeny of New Zealand Cyanidiales inferred from maximum likehood (ML) analysis of the combined DNA sequences of *psa*A, *psb*A, *rbc*L, 18S rRNA and plastid 16SrRNA genes. Values from Bootstrap analysis are shown on above the branches as decimals. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets

4.3.4 Plastid 16S rRNA gene BLAST search

Table 4.3 Nearest BLAST relatives of CM1.2Y, CM1.2G and RK1.2

Isolate	Accession	Organism description	Pairwise identity (%)
CM1.2Y	X52985	*C. caldarium* plastid 16S rRNA	99.9
	KJ00459	*G. sulphuraria* 074W complete plastid	99.7
	KJ00459	*G. sulphuraria* 074W complete plastid	99.7
	AF170718	*G. sulphuraria* SAG107.79 16SrRNA	99.7
	AF356019	Rhodophyte NZ4 plastid 16S rRNA	99.2
CM1.2G	AY391361	*G. maxima* IPPAS P507	99.8
	KJ569775	Cyanidiceae sp. MX-AZ01 complete plastid	99.8
	MK231134	*C. merolae* isolate 5508	96.9
	AY286123	*C. merolae* DBV201 complete plastid	95.8
	AF545617	*C. merolae* DBV201 16S rRNA plastid	95.8
RK1.2	KJ00459	*G. sulphuraria* 074W complete plastid	99.4
	KJ00459	*G. sulphuraria* 074W complete plastid	99.4
	X52985	*C. caldarium* plastid 16S rRNA	99.3
	AF170718	*G. sulphuraria* SAG107.79 plastid 16SrRNA	99.1
	AF356019	Rhodophyte NZ4 plastid 16S rRNA	98.9

The plastid 16S rRNA gene Blast results are similar within the yellow (excluding RK2) isolates and green isolate groups, with low variance in the pairwise identity percentage (≤0.2%). So CM1.2Y, CM1.2G and RK1.2 results are shown as representative of the different groups.
The BLAST results show that plastid 16S rRNA sequence of CM1.2G isolate show a high similarity with a *G. maxima* IPPAS P507 (99.8% Pairwise identity) and several *C. merolae* strains (Table 4.3). This is consistent with the other phylogenetic data in this thesis.
The BLAST results also show that the plastid 16S rRNA gene of both CM1.2Y and RK1.2 isolates share a high similarity with a *C.caldurium* plastid 16S rRNA gene (99.9% and 99.3% pairwise identity respectively table 4.3). Both CM1.2Y and RK1.2 isolates have high similarity to "Rhodophyte NZ4", which is the Cyanidiales previously isolated in White Island, NZ [27]. This is in conflict with other phylogenetic data in this thesis, which indicates a relation to *G. sulphuraria* strains. However, the five results with the highest grade include; the two copies of 16S rRNA gene from the *G. sulphuraria* plastid genome and a plastid 16S rRNA gene from *G. sulphuraria* SAG 107.79, for all yellow isolates. That is consistent with the other phylogenetic data in this thesis.

4.4 Discussion

4.4.1 Phylogeny and classification of isolates

The molecular and phylogenetic data gathered in this thesis supports the hypotheses that;
a) New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* lineage and b) these lineages can be differentiated using phylogenetic analysis. All the green isolates are very closely related and are *G. maxima* related strains. As in all the phylogenetic trees, they cluster together and form a monophyletic clade with other known *G. maxima* strains, including the strain IPPAS P507 which is the type strain for *G. maxima*[61]. The green isolates are also closely related to the previously isolated NZ type V strains, given the phylogeny of the 18S rRNA and *rbc*L gene sequence (fig 4.2 and 4.4). None of the isolates from this thesis had a close relation to the NZ type VI strain.
The yellow strains are all also *G. sulphuraria* strains, given the consensus of all the phylogenetic trees. However, not all of the yellow isolates share the same lineage; isolates RK1.3, CM1.2Y, CM1.3Y, CM1.355 and KP1.2Y are all very closely related and share a common lineage with *G. sulphuraria* strains isolated from YNP. Meanwhile isolate RK1.2 is distinct, on the phylogenetic

trees it is separate from the other yellow isolates, forming a monophyletic clade with *G. sulphuraria* strains isolated from Italy.

The original discovery, which this data confirms, that the New Zealand Cyanidiales population contains *G. sulphuraria* strains with a shared lineage with American strains and *G. maxima* strains with a shared lineage with Eurasian strains, was impactful as it was evidence of two separate migration events from different continents. This had not been observed at most sites that had both *G. maxima* and *G. sulphuraria* strains. For example, in Iceland, Japan and Taiwan, the *G. maxima* and *G. sulphuraria* both shared a Eurasian lineage, all being closely related to Russian strains and to each other [32,38,78,161].

The phylogeny of RK1.2 is evidence of a third migration event from a European lineage and potentially unsampled strains and sites between New Zealand and Italy.

There is evidence, from other phylogenetic studies, that the Italian *G. sulphuraria* strains are more closely related to American strains than to other Eurasian strains [38,77,78]. Additionally, a recent phylogenetic analysis of the *rbc*L gene from 81 novel Cyanidiales isolates from Turkey detected all known species of Cyanidiales, including both *G. sulphuraria* from the Eurasian lineage and from the Italian/American lineage [159].

Therefore it is possible that the phylogenetic data might be exaggerating the number of migration events of Cyanidiales to NZ. There could have been a single migration of *G. sulphuraria* and the divergence occurred in NZ. There is the possibility that multiple lineages were established in a single migration event, given that the lineages can coexist in very close proximity and that there could be more sites with high Cyanidiales diversity like Turkey, that are not yet sampled. To accurately determine lineages to this degree would require the phylogenetic analysis of a much larger and more diverse sequence dataset than used in this thesis, and would be out of the scope for this thesis.

It can be concluded that the green isolates (CM1.1G, CM1.2G, CM1.3G and KP1.2G), the yellow isolates (CM1.2Y, CM1.3Y, CM1.355, KP1.2Y and RK1.3) and RK1.2 are phylogenetically distinct, representing one *G. maxima* strain and two separate *G. sulphuraria* strains. It was not possible to distinguish RK1.2 from the other yellow isolates using visible morphology, as the cells and colony morphology were identical (figs4.10). If it had been chosen for further characterisation, it would be possible that no significant differences would be observed given how difficult it was to differentiate between *G. sulphuraria* strains using morphological, eco-physiological and biochemical methods [36,77].

It is difficult to determine the relationship of *G. sulphuraria* isolates to the previously isolated NZ type IV strains. The cell morphology described (fig 2.5 G) is similar to the yellow isolates and *rbc*L phylogeny does indicate a shared America lineage. However, the 18S rRNA gene phylogeny of the yellow isolates does not indicate high homology to *G. maxima*, as observed with NZ type IV. There is evidence that this conflict of phylogeny in the NZ type IV is due to a mixed culture and primer bias (discussed in section 4.4.2), and that *rbc*L is the best indicator of the true phylogeny of NZ type IV strains.

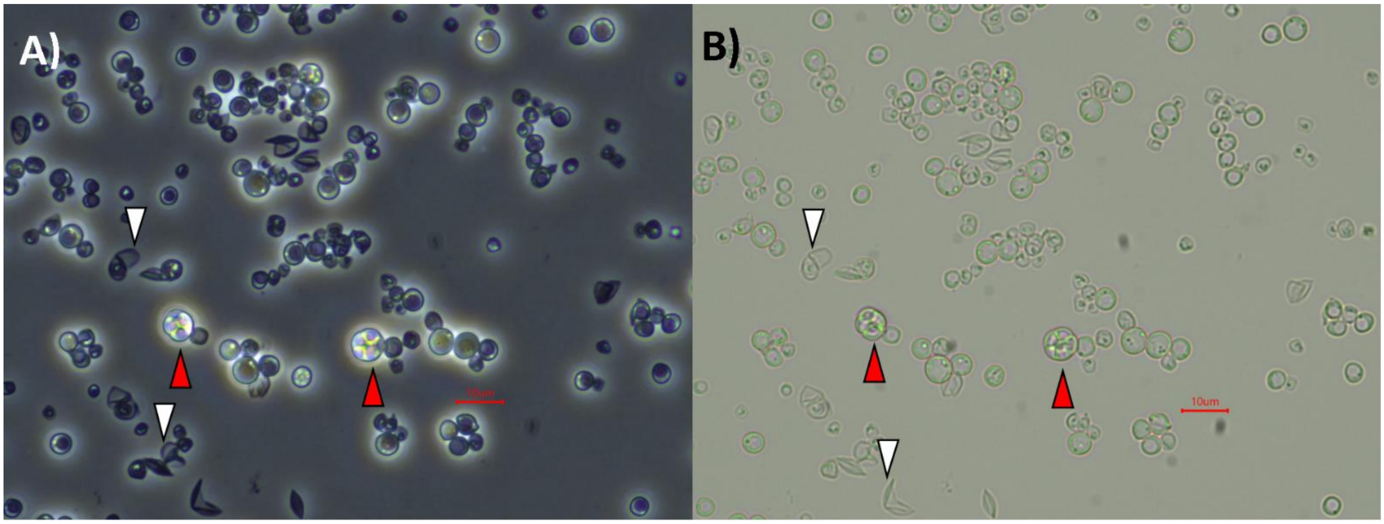


Figure 4.10: Microscope images from axenic cultures of the yellow isolate **RK1.2** as observed using phase-contrast(A) and bright field (B) with 40x objective. Red triangles indicate vegetative cells with endospores. White triangles indicate cell wall debris from lysed vegetative cells. Scale bar is 10µm

4.4.2 The issues with 18S rRNA sequencing

The 18s rRNA primer sequences used in this thesis and in Toplin *et al.* 2008 were originally from Gross *et al.* 2001. As the aim was to determine the relation of the isolates from this thesis to Cyanidiales previously isolated from NZ, it was decided to use the same 18s rRNA and *rbc*L primers as Toplin *et al.* 2008 to determine if there were novel strains isolated,. However, there were issues experienced with the 18S primers. Initially, there were issues with amplification with PCR. Which once remedied there were issues with sequencing. Figure 4.11 show the chromatograms of CM1.3Y and CM1.1G as viewed in Geneious. The chromatograms of CM1.3Y indicate acceptable base call quality of the sequencing of the forward direction, but poor quality in the reverse direction. While chromatograms of CM1.1G show good base call quality in both directions. This pattern was repeated in all green isolates and in all yellow isolates. After viewing the chromatogram of the initial sequences, a re-sequencing request was sent to Macrogen. However, the resequencing showed the same results. So new PCR products were generated from new extractions and with fresh aliquots of primers. These were then resent to Macrogen for sequencing, which were sequenced twice, resulting in four sequences for each of the forward and reverse directions. While the reverse reads were poor for the yellow isolates, given the three to four replicates of the forward direction, there was high confidence in the sequence quality. If the yellow cultures contain cells from the green isolates the reverse sequence would likely appear to amplify correctly, however the consensus sequence generated would be incorrect and have homology close to *G. maxima*. There is other evidence that issues occurred in Toplin *et al.* 2008. There are 59 *rbc*L gene sequences published in Toplin *et al.* 2008 but only 39 18S rRNA sequences published. The three 18S rRNA sequences published from NZ type IV strains; CCMEE 5706CR, 5711WI and 5720WI were 468bp, 242bp and 439bp in length respectively. The primer set is designed to generate sequence lengths up to 750bp, in this thesis sequence lengths were 600-700bp for green isolates and 500-600bp yellow isolates after trimming. Additionally, Toplin *et al.* 2008 used serial dilution to extinction method to isolate strains using an autotrophic medium. Therefore, the different morphologies would not have been observed.

There are two conclusions here. First, that the primer set from Gross *et al.* 2001 is not suitable for all Cyanidiales strains and a new 18S rRNA primer set needs to be developed.
Second the importance of heterotrophic media in the accurate isolation of Cyanidiales strains. The history of Cyanidiales is full of mixed cultures that have been separated by use of heterotroph medium. The original *Cyanidium caldarium* culture was separated into forma A and forma B following observations of different morphologies when grown in the dark [48]. The original Russian strains *G. maxima, G. daedala* and *G. partita* were isolated on heterotrophic media from a mixed culture [61]. Heterotrophic growth was distinguished between two *G. sulphuraria* strains from the same sample due to differences in pigment profile [91]. Isolation on heterotrophic should become a standard methodology when isolating Cyanidiales in the future.

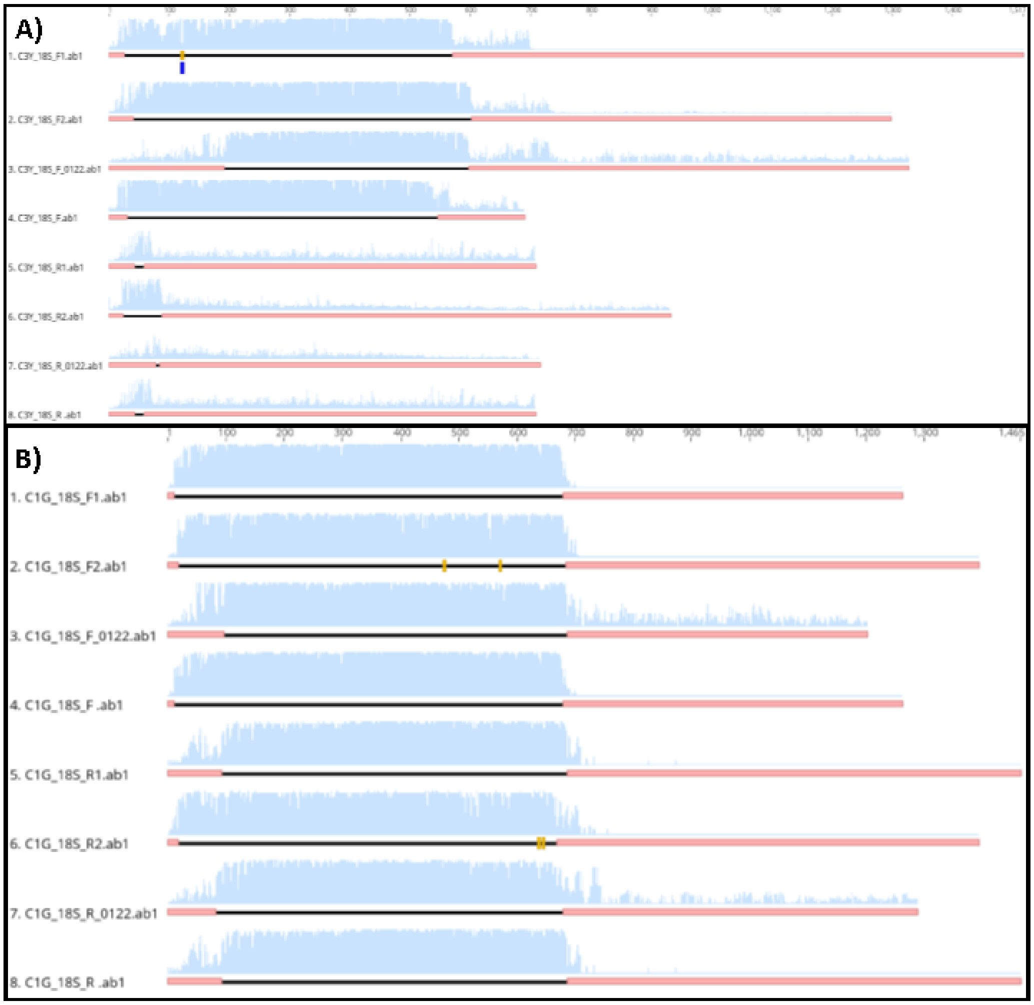


Figure 4.11: Chromatograms of single direction sequencing reads as viewed in Geneious. Pink bars indicate trimmed regions. Blue indicate base call quality A) Is the four forward and four reverse reads from the Yellow isolate CM1.3Y B) Is the four forward and four reverse reads from the Green isolate CM1.1G

4.4.3 Relation to Cyanidium caldarium

The plastid 16S rRNA BLAST indicates that the yellow isolates have high similarity to *C. caldarium* (99.3-99.9%), which is in conflict with the other phylogenetic data and the other BLAST results which include other *G. sulphuraria* strains. The top five BLAST results for Green isolates include *G. maxima* IPPAS P507 (99.9%) and *Cyanidioschyzon* strains (96-97%), but not any sequences from *Cyanidium* strains. This indicates that the plastid 16S rRNA is conserved to a degree within a Family unit and between sister genera, however not to the degree that would explain a 99% similarity between *G. sulphuraria* and *C. caldarium*. This is the consequence of the misidentification and misnaming of a *G. sulphuraria* strain. Tracing the accession (X52985) of the *C. caldarium* sequence indicates it is from the *C. caldarium* strain SAG107.79 [171]. This strain was originally isolated YNP [48] and was later reclassified as *G. sulphuraria* SAG107.79 [36]. The plastid 16S rRNA gene sequence was republished later, accession number AF170718 [172] which the yellow isolates also have high similarity (99.1-99.7%) to the republished sequence (Table 4.3). At the time that Donachie *et al.* (2002)[27] isolated Rhodophyte NZ4, there would have been very few *G. sulphuraria* plastid 16S rRNA sequences published, likely leading to the misidentification. It is possible the same error occurred in the 16S rRNA survey of Norris basin, YNP, in which *C. caldarium* was identified as the primary Eukaryote [83], but no accession numbers are given in the paper [83].

It can be concluded that Rhodophyte NZ4 was a *Galdieria* strain, however it appears that neither RK1.2 or the other yellow isolates are the same strain. Similarity of (98.9-99.2%) for the plastid 16S rRNA gene indicates a significant degree of divergence, so the diversity of Cyanidiales may be even greater than indicated by the sampling in this thesis and in Toplin *et al.* 2008.

4.5 Conclusion

Phylogenetic analysis indicates that the ten isolates, collected in this thesis, represent strains from the *G. maxima* lineage and two *G. sulphuraria* lineages. This supports the hypotheses that New Zealand Cyanidiales strains descended from at least two lineages, the *G. maxima* lineage and the *G. sulphuraria* and that these lineages can differentiated using phylogenetic analysis. The *G. sulphuraria* isolates plastid 16S rRNA gene did show high similarity to a *C. caldarium*, although that was due to the *C. caldarium* sequence being misnamed. The RK1.2 represents a novel lineage not observed in New Zealand before and was not detected by observations of visible morphology.

There is evidence that the NZ Cyanidiales population is even more diverse, given that no clear examples of NZ type VI or Rhodophyte NZ4 were isolated in this thesis. It indicates that sampling in NZ is far from saturation, and other strains can be sampled and isolated in future work.

Chapter 5

5.1 Introduction

Aims of this the thesis were;

1. Sample geothermal sites in New Zealand for Cyanidiales.
2. Enrich for and isolate Cyanidiales strains.
3. Carry out basic characterisation and phylogenetic analysis to determine taxonomy.
4. Assess the individual isolates suitability for use in the theoretical bioreactor.

Aims one to three were done empirically and are detailed in Chapter 3 and 4. Which provides evidence to support the hypotheses that; A) NZ Cyanidiales descended from at least two lineages; the *G. maxima* and *G. sulphuraria* lineages, and B) that these lineages can be differentiated by phenotypic characterisation and phylogenetic analysis. Additionally, the phenotypic and phylogenetic data indicate that the NZ *G. maxima* are atypical to the type strain, therefore are significant in the understanding of the evolution of *G. maxima* and Cyanidiales. A summary of Chapter 3 and 4 are given in this section, as well as the evidence that support these hypotheses.

Aim four is related to the background of this thesis (section x.x), the development of a photobioreactor that will use Cyanidiales strains to sequester carbon emissions from flue gas to generate biofeedstock. Given the scope and time frame of a Master's thesis this cannot be done empirically. The assessment of the NZ Cyanidiales strains suitability for use in the theoretical bioreactor is a desktop assessment. Using the data from characterisation and from the literature of related strains, I can offer advice as to which of the strains isolated in this thesis are most suitable for use in the proposed photobioreactor.

5.2 Assessment of NZ Cyanidiales suitability for use in the theoretical bioreactor

5.2.1 Introduction

My role, as part of this thesis, is to contribute New Zealand Cyanidiales strains to a larger research of a proposed photobioreactor, that will capture waste carbon gas emissions to generate biomass as biofeedstock (section x.x). The basic concept is to capture waste carbon gas emissions from a point source, such as a geothermal power plant which will consist of mostly CO_2 and CH_4 and feed the gas through a photobioreactor. The photobioreactor will contain a consortium of extremophilic methanotrophs and photoautotrophs microorganisms. The role of the methanotrophic organisms will be to oxidise the methane and other short chain hydrocarbons, generating CO_2 and biomass as growth. The photoautotroph converts the CO_2 to carbohydrates, biomass and O_2. The additional O_2 supports the growth of the methanotrophs. Cyanidiales were proposed as the extremophilic photoautotroph as optimal growth conditions overlapped with those of the candidate methanotrophs, e.g. *Methylacidiphilum* V4[23,173] (section x.x), and are native to New Zealand. However as discussed (section x.x), there were a limited number of NZ Cyanidiales that had been isolated and characterised.

One of the primary goals was to provide this project with additional characterised native New Zealand Cyanidiales strains to ensure successful development of the project. In this thesis, ten NZ Cyanidiales have been isolated. Of these, four isolates (CM1.1G, CM1.2G, CM1.3G, KP1.2G) are from the *G. maxima* lineage, five isolates (CM1.2Y, CM1.3Y, CM1.355 KP1.2Y RK3) are from an American *G. sulphuraria* (refer to as *G. sulphuraria* (A) in this section) lineage and a single isolate (RK1.2) is from an Italian *G. sulphuraria* lineage. There is very little observed variation within the *G. maxima* lineages and *G. sulphuraria* (A) lineages, based on the characterisation of two isolates from each of those lineages(Chapter 3). The isolate RK1.2 was not characterised to the same degree as the other lineages, however given the similarities between *Galdieria* species and strains [36,77], it can be assumed that RK1.2 will have similar characteristic to the *G. sulphuraria* (A) lineage for the purposes of assessing the application to this proposed bioreactor.

5.2.2 Assessment of NZ Cyanidiales in consortium with Methanotrophs

Research and development of Cyanidiales biotechnology typically uses *G. sulphuraria* strains [47] due to the species' flexible metabolism [47,91] and high tolerance to stress factors, such as heavy metals [44,46] or salts [31,36]. This has allowed for high density *G. sulphuraria* cultures to be grown with a wide range of substrates, including bakery waste [115] and primary effluent [21,123], for wide range of functions such as production of C-phycocyanin [106,113–115], single-stage treatment of urban wastewater [21,122,123,174], remediation of heavy metal pollution [119,120] and recovery of valuable metals [116–118].

Evidence from this thesis shows that the *G. sulphuraria* (Y) isolates remain viable up to 28 days at 55°C, are facultative heterotrophs and growth was observed to pH 0.5. Therefore, these strains will be capable of growth in conditions suitable for *Methylacidiphilum* V4 [22,23] and of surviving potential fluctuations in the temperature and pH that can occur in bioreactors. However, for the proposed bioreactor, *G. sulphuraria* strains may not be the optimal choice for several reasons. The primary purpose of the Cyanidiales strain in the proposed bioreactor is to convert CO_2 to biomass, through photoautotrophic growth. Therefore, a facultative heterotroph is not required and could potentially reduce the efficiency of the bioreactor. Both the RK1.2 and the other *G. sulphuraria* isolate reduce production of the photosynthetic pigments under heterotrophic growth conditions, resulting in a long-lag phase when returned to autotrophic growth conditions [91,113]. A lag-phase of over 28 days was observed for RK1.3 isolate. There are *G. sulphuraria* that maintain pigment production during heterotrophic growth [91,113], but this phenotype was not observed in any of the isolates from this thesis. There is the potential that conditions in the bioreactor may induce heterotrophic growth and induce the reduced pigment phenotype, due to alternative carbon sources becoming available, e.g. the breakdown of older cells. This would reduce the efficiency of the bioreactor, as the rate of O_2 production from Cyanidiales biomass would decrease and total CO_2 would increase, which would affect the rate of methane oxidation by the methanotrophs. So ideally in the proposed bioreactor, the Cyanidiales strain needs to be an obligate photoautotroph or a *G. sulphuraria* that does not lose pigment during heterotrophic growth.

Another concern is the competition for trace metals. The ability of *G. sulphuraria* strains to selectivity uptake and bioaccumulate rare and valuable metals is well documented [116–118,120]. Over 70% of rare metal ions, such as neodymium (Nd), lanthanum (La) and dysprosium (Dy), were recovered from solution within 24 hours by *G. sulphuraria* 074W[117]. The theory is that the primary purpose of this bioaccumulation is to reduce oxidative stress from heavy metals [44,46,116]. This would be useful in biotechnology applications such as electronic recycling or acid-mine drainage remediation but is potentially detrimental to the proposed bioreactor.

The methanol dehydrogenase (MDH) of acidophilic methanotrophs, such as *Methylacidiphilium* spp., is dependent on rare earth metals such as La, cerium(Ce) and Nd [135]. The MDH enzyme is key in the methane oxidation pathway, as such the V4 medium [23,135] contains added La and Ce(section x.x). There is potential that the Cyanidiales strain used will take up and bioaccumulate these rare earth metals, reducing the growth of the methanotroph, rate of methane oxidation and overall efficiency of the bioreactor. This could also affect the quality of the final feedstock product as this will concentrate other toxic heavy metals within the biomass, making it unsafe for consumption. In vitro studies have only been done using *G. sulphuraria* strains. However, from genome research there is evidence that all Cyanidiales have enzymes to reduce heavy metal toxicity, e.g. mercuric reductase, but *C.merole* strains lack the variety of metal transporter found in *G. sulphuraria* strains [44,46,57]. Without these metal transporters, *C. merolae* strains will be incapable of uptaking these rare metal ions. Given that the closest relative of *G. maxima* is *C. merolae*, it is probable that *G. maxima* lacks many of these metal transporters as well. However, there is no genome data or experiments of the uptake of metal ions using *G. maxima*. As such, experiments may need to be done or the genome can be searched for metal transporters genes.

Another potential issue with the *G. sulphuraria* (Y) isolates is that the strain is only able to utilize ammonium as an inorganic nitrogen source. While the concept of balancing nitrate and ammonium uptake as a metabolic pH control is in the early stages of research and would require experiments with the specific strain[144,145], it is not possible with *G. sulphuraria* isolates. Although the nitrogen source utilization of RK1.2 still needs to be tested.

Given these reasons, I would recommend the use of a *G. maxima* isolate (CM1.1G, CM1.2G CM1.3G, KP1.2G) if the proposed bioreactor requires the methanotroph and photoautotroph to be cultured in consortium. Growth of CM1.3G and KP1.2G was observed up to 50°C and to pH 1.0 (section x.x), therefore these isolares can grow in conditions suitable for *Methylacidiphilum* V4. Additionally, there is evidence that these isolates have a limited capability to metabolise organic compounds which enhances growth rates but does not result in an observable change in pigment. The rare metal uptake and accumulation abilities are unknown at this point, however once the genome is sequenced and annotated, it can be searched for metal transporters protein-coding sequences.

5.2.3 Summary of assessment

In summary, for the proposed photobioreactor I would recommend the use of either the KP1.2G or CM1.3G strains. The temperature and pH growth ranges of these strains overlap with those of the candidate methanotrophs. Therefore, conditions within the bioreactor can be maintained to optimise the growth of both the methanotrophs and the Cyanidiales strain. Additionally, as these strains are obligate photoautotrophs this means the rate of CO_2 fixation will be predictable. While no empirical data is available yet, genome data from related strains indicates these strains lack metal transporters and so will not starve the methanotrophs of the essential rare metals.

5.3 Summary of research and Final discussion

In this thesis, Chapter 3 detailed the sampling and isolation of Cyanidiales from the Taupō Volcanic Zone (TVZ) and the phenotypic characterization of a selection of those Cyanidiales. Chapter 4 detailed the collection of molecular data from the Cyanidiales sampled and the phylogenetic analysis of that data.

The combination of data from Chapters 3 and 4 support the hypotheses that NZ Cyanidiales descended from at least two lineages; the *G. maxima* and *G. sulphuraria* lineages, and that these lineages can be differentiated by phenotypic characterisation and phylogenetic analysis. Additionally, the NZ *G. maxima* are phenotypically atypical, and present an opportunity to determine the evolutionary pathway of *G. maxima* and revise the nomenclature.

5.3.1 Summary of Chapter 3

In Chapter 3, the ten Cyanidiales representatives were isolated from New Zealand and were separated into two groups, the Yellow isolates and the Green isolates, based on the colony morphology on MAll-HT plates and cell morphology. The growth range, capability to utilize nitrogen sources, heterogrowthic growth in the dark and lipid profile was determined for two green isolates, CM1.3G and KP1.2G, and two yellow isolates, RK1.3 and CM1.3Y.

Isolates RK1.3 and CM1.3Y, from the Yellow group, were characterised and described *G. sulphuraria*. There was no observable difference between these two strains, despite the geographic distance between the sampling sites. Additionally, there were no differences observed in cell morphology of the other yellow strains. This indicated that all the yellow isolates were from the *G. sulphuraria* lineage. Isolates KP1.2G and CM1.3G, from the Green group, were characterised and described as representatives of the Cyanidiaceae family. However, it was difficult to describe the genus of these green isolates as the club shape cells were indicative of *C.merole* (fig 2.1), but there was evidence of endospores in KP1.2G and CM1.3G (fig 3.6 3.7). Therefore, the reproductive pattern indicated a closer relation to *Cyanidium* than *Cyanidioschyzon*. An additional complication was that the cell morphology was similar to the description of the previously isolated NZ type V strain (fig 2.5 h), a strain that was related to *G. maxima* based on the phylogeny or the *rbc*L and 18S rRNA gene [26]. The type strain of *G. maxima* P507, is a facultative heterotroph with a cell morphology indistinguishable to other *Galdieria* (fig 2.2 [36,61]), meanwhile testing indicated that the green isolates were incapable of heterotrophic growth in the dark. While the phenotypic characterisation of the isolates did differentiate between the yellow and green isolates, there was not enough resolution to determine taxonomy to the species level, and determining the taxonomy required molecular and phylogenetic data.

5.3.2 Summary of Chapter 4

In Chapter 4, a set of five genes were sequenced from each of the ten isolates. The genes were the 18S rRNA, plastid 16S rRNA, *psa*A, *psb*A and *rbc*L genes. Phylogenetic analysis was carried out on single-gene datasets for *rbc*L and 18S rRNA genes and concatenated dataset, which included a three plastid protein gene (*psa*A, *psb*A and *rbc*L) dataset and a five-gene dataset.

The *rbc*L and 18S rRNA trees (figs 2.2 2.4) did indicate that the green isolates were related to the previously isolated NZ type V strains and *G. maxima* P507. While the majority of the yellow isolates, excluding RK1.2, had a common lineage with the previously isolated NZ type IV, YNP type II strains [26], and described American *G. sulphuraria* strains such as SAG108.79 [36,48]. The single RK1.2 isolate was within the *G. sulphuraria* clade, but clustered with described Italian *G. sulphuraria* strains [38] away from the other isolates from this thesis or previously isolated NZ strains.

The phylogenetic analysis of the concatenated datasets (figs 2.5-9.) were consistent with the single-gene trees. All the green isolates were closely related, representing a single strain of the *G. maxima* lineage and are examples of the previously isolated NZ type V strains.

All the yellow isolates are from the *G. sulphuraria* lineage; however these yellow isolates represent two strains with distinct lineages. The RK1.2 represents an Italian *G. sulphuraria* lineage, not previously documented in NZ. All other yellow isolates represent an American *G. sulphuraria* lineage that includes previously isolated NZ type IV strains and other *G. sulphuraria* strains from YNP.

It is difficult to determine the relationship between the American *G. sulphuraria* isolate from this thesis and the NZ type IV strains from Toplin *et al.* 2008, due to the conflict in the phylogeny of *rbc*L and 18S rRNA genes for NZ type IV strains. The cell morphology and *rbc*L gene phylogeny indicate NZ type IV strains represent *G. sulphuraria,* while the phylogeny of the 18S rRNA gene indicates a relation to the *G. maxima* lineages. This conflict was not observed in the *G. sulphuraria* isolates from this thesis. It is possible that a “true” representative of NZ type IV was not sampled within this thesis, however sampling sites from which NZ type IV strains were isolated include Craters of the Moon, a sampling site chosen for this thesis. I provide evidence (section 4.4.2) for the hypothesis that the conflict between *rbc*L and 18S rRNA gene phylogeny is due to a non-axenic NZ type IV cultures and that the 18S rRNA gene primers are biased towards amplification of sequences from Cyanidiaceae species. Although, this would be difficult to determine conclusively without obtaining the cultures from the CCMEE, but is a potential avenue of future research. Based on the evidence from the *rbc*L phylogeny (fig 4.2), it can be determined that NZ type IV and the *G. sulphuraria* isolates (excluding RK1.2) have a common lineage.

5.3.3 New Zealand Galdieria Maxima and resolving the evolution of the *G. maxima*/Cyanidioschyzon lineage

Taken together, the morphological, ecophysiological, biochemical and genetic data indicates that the three distinct Cyanidiales strains were isolated from NZ in this thesis, which includes a single NZ *G. maxima* lineage and two *G. sulphuraria* lineages, one American and one Italian (RK1.2) lineage. This supports the hypotheses that NZ Cyanidiales descended from at least two lineages, the *G. maxima* and *G. sulphuraria* lineages, and that these lineages can be differentiated by phenotypic characterisation and phylogenetic analysis. The RK1.2 Italian lineage is significant as it had not previously been observed in NZ. While the *G. maxima* lineage had been previously isolated (NZ type V strains [26]), is also significant as the usual phenotypic characteristics compared to the *G. maxima* type strain, P507, had not been observed. The obligate autotroph, "*Cyanidium*-like" NZ *G. maxima* presents a potential opportunity to determine the evolutionary pathway between *C. caldarium* and the *Galdieria*-like *G. maxima* P507, and could lead to a revision of the nomenclature and amendments to the description of *G. maxima*. Similarly, Toplin *et al.* 2008 isolated YNP type IA strain, a *C. merolae* related strain that has a cell wall, which was noted as potentially representative of the evolutionary pathway between the cell walled Cyanidiaceae ancestor and the cell wall-less *C. merolae* ecotypes. These strains together are evidence that the common ancestor of the *G. maxima* and *Cyanidioschyzon* was *C. caldarium*-like, i.e. an acido-thermophillic photoautotroph with a cell wall. It is likely that the *Cyanidioschyzon* and *G. maxima* lineages diverged fairly recently, after the divergence from the *Cyanidium* lineage approximately 350Mya [74]. The *Cyanidioschyzon* ecotype is due to the common ancestor inhabiting osmotically stable environments and the rapid genome reduction[76] with the loss of non-critical pathways such as the floridean starch pathway, and consequently the ability to produce a cell wall[42,43,57,95]. Floridean starch was not critical as, unlike other Rhodophyta, *Cyanidioschyzon* strains did not require the floridean starch as an osmolyte [93,94,175], and the primary carbon storage compound for all Cyanidiales was glycogen-like polyglucans [70,175–177]. The *G. maxima* ancestor diverged from the *Cyanidioschyzon* lineage and acquired carbon transporters genes through HGT events, displaying convergent evolution with the *Galdieria* lineage and allowing the colonization of a greater variety of environments [44,46].

These theories are primarily based on phylogenetic and morphological evidence, the close relation of the *G. maxima*/*Cyanidioschyzon* to *C. caldarium* [38] and the shared morphology between the lineages, and evidence for genome reduction [76] and HGT in Cyanidiales from data from the several *Galdieria* and two *C. merolae* genomes [42,44,46]. However, the genome sequence from Cyanidiaceae representatives is limited to the two *C. merolae* strains, with no genomes from either *G. maxima* or *Cyanidium* strains. The genome sequences of a YNP IA strain, an NZ *G. maxima* strain, *G. maxima* P507 and a *C. caldarium* strain would provide substantial evidence to determine the validity of these theories[99]. This could be an avenue for future research, however this is ambitious and would require the collaboration of many researchers over potentially many years of work. Sequencing the genome of KP1.2G is more feasible and would still contribute to the understanding of the evolution of *G. maxima,* leading to the revision of the nomenclature.

5.3.4

The observations and results from this thesis stress the importance of using both phenotypic characterization and phylogenetic analyses in the isolation and characterisation of Cyanidiales strains. The phenotypic characterisation in this thesis lacked the resolution to accurately determine the taxonomy of the NZ *G. maxima* isolates past the family level and RK1.2 could not have been identified as a distinct *G. sulphuraria* lineage compared to the other NZ *G. sulphuraria* isolates. While only molecular data and phylogenetic analysis would have been sufficient to accurately determine the taxonomy of these NZ strains, the usual and significant phenotypes of the NZ *G. maxima* would not have been observed, whichas had occurred when NZ type V strains had previously been isolated[26]. *Galdieria maxima* strains have been isolated from a number of other sites including Japan, Iceland and Turkey [26,32,159], however the isolation, in all cases, used autotrophic media and taxonomy was determined primarily with phylogenetic analysis and very limited phenotypic characterisation [26,32,159]. As such, the prevalence of facultative heterotrophs in the *G. maxima* lineage is unknown, therefore it is possible that it is limited to *G. maxima* IPPAS P507. Likewise, many of these cultures may not be axenic cultures, given that the historical isolation of both *G. sulphuraria* [30,53] and *G. maxima* [61] were from mixed "*Cyanidium caldarium*" cultures using media supplemented with glucose [30,50,55]. Isolation using a medium supplemented with glycerol and glucose, as used in this thesis, enhances growth while allowing and allows for observation of pigment changes and heterotrophic growth. If bacterial contamination is a concern, there is evidence from both this thesis and the literature [136] that a number of antibiotics such as ampicillin and vancomycin do not inhibit the growth of Cyanidiales. If there are still concerns, with isolation on solid heterotrophic media can be used as a secondary test after the primary isolation. There could be concerns for the isolation of osmotically sensitive, strict photoautotroph strains such as the *C. merolae* ecotype, however the Modified Allen's medium recipe was optimised specifically for culturing *C. merolae* and enhanced growth was observed when 0.4% glycerol was added [136], as observed in this thesis. Therefore, I would recommend that for all future isolation of novel Cyanidiales strains, the Streak plate technique on solid medium, supplemented with glycerol and glucose, should be preferred over serial dilution to extinction using autotrophic medium.

5.4 Future work

With the isolations and characterisation of three NZ Cyanidiales strains, a number of avenues for future research have been opened up.

The most realizable avenue is to sequence the genome of KP1.2G and RK1.3, using the prepared extracts (section x.x). There are 11 *Galdieria* genomes published at this time [44–46] so the genome of RK1.3 would be used to detected novel genotype, suitability to other biotechnologies and determine the taxonomy and the relation of NZ *G. sulphuraria* to American strains with higher resolution. The sequencing of the KP1.2G nuclear genome is more significant, as at the time of writing, it would represent the first *G. maxima* genome sequenced and the third from the Cyanidiaceae family[46]. The other two Cyanidiaceae genomes are from*C. merolae* 10D [42,43] and Soos [46]. Ideally, a culture of *G. maxima* P507 which is the facultative heterotroph type strain, would be acquired from IPPAS meaning the genome could be extracted and sequenced as well. This would give the data required to determine the evolutionary pathway between *C. merolae*, an autotroph that lacks a cell wall, and *G. maxima* P507, a facultative heterotroph that has a cell wall. With KP1.2G as a likely stepping stone between the lineages. If the genomes of the *G. maxima* strains are more closely related to *Cyanidioschyzon* genomes, it would be strong evidence to resolve the taxonomy leading to a revision of the nomenclature and descriptions. I would hypothesise that *G. maxima* P507 acquired the capabilities to transport carbon sources through a recent HGT, independent of the true *Galdieria* lineage. Which could indicate that it is a poor type strain and that most *G. maxima* strains and species of the Cyanidiaceae family are obligate autotrophs.

Besides resolving taxonomy, just the genome sequence of KP1.2G could resolve a number of inquiries as well. The NZ *G. maxima* strains are capable of utilising urea as the sole nitrogen source. As discussed previously (section 3.x.x), most Cyanidiales have lost the gene-set required for uptake and assimilation of urea with the known expectation of *G.phlegrea* DBV009[45]. This strain acquired a full functional set of urease genes through multiple HGTs from Eubacteria, but are geographically and phylogenetically (divergence 700-900 Mya [46]) distant to NZ *G. maxima* strains. As such, the questions are: does the NZ *G. maxima* have a full urease gene-set and are the origins eukaryotic, i.e. derived from the Archaeplastida or Rhodophyta last common ancestor, or prokaryotic, i.e. derived from an HGT event. Given that the *C.merole* 10D lacks all urease genes [42,45] and is the sister lineage to the *G. maxima* lineage[32,38], I would hypothesise that KP1.2G genome will have the full urease gene set which was acquired by HGT events, independent of *G.phlegrea* DBV009. There is the possibility that the urease gene-set would be partially or completely Eukaryotic in origin or there is an alternative urea assimilation pathway.

The genome could also be searched for carbon or metal transporter proteins. The presence of carbon transporter could indicate that the *G. maxima* strains are capable of heterotrophic growth, however it might require a different carbon source besides glucose or glycerol. The presence of metal transporters will indicate if these NZ *G. maxima* strains are capable of starving the methanotroph of rare earth metals. I would hypothesise that the KP1.2G genome, like *C. merolae* genomes [44,57], will lack metal transporters and metal tolerance will be due to metal detoxifying enzymes like mercuric reductase [44,46].

Beside the genomic research, there are a number of additional experiments and experimental work to be retested.

The strain RK1.2 needs to be characterised, as described in Chapter 3, to determine a description and detect any novel phenotypes.

Ideally, the pH growth range experiment needs to be redone with a medium that has a high concentration of buffers, or in chemostat cultures to determine the upper pH growth limit of these NZ strains.

Similarly, with the NZ *G. maxima* reduced growth was observed on nitrate compared to ammonium. This needs to be retested in medium with a higher concentration of buffers to determine if the reduced growth was due to the alkalinization of the medium.

Further research could be in the development of the proposed bioreactor, determining empirically which Cyanidiales strains are most suitable in consortium with the candidate methanotroph strains. This could be determined using several different measurements such as highest growth rates, carbon gas conversion efficiency and/or greatest nutritional value.

There is also the potential of further sampling and isolation of novel Cyanidiales in New Zealand. The sampling within NZ is still limited, but even so, the sampling in this thesis isolated a novel strain (RK1.2) not detected previously in NZ [26,27]. Although did not isolate a clear representative of the NZ type VI strain [26] or of Rhodophyte NZ4 [27], as such these strains remain uncharacterised. The NZ type VI strain includes a single isolate from Waiotapu, so it might not be as widespread as the *G. maxima* (V) or *G. sulphuraria* (IV) lineages, but the phylogenetic analysis indicates it is a *G. maxima* strain as well[26]. However, the cell morphology of NZ type VI with a large round cell (fig. 2.5, I) is closer to cell morphology of *G. maxima* P507 (fig. 2.2.) so may indicate it is closer phenotypically to the *G. maxima* P507[36]. The NZ type VI strain might be limited to Waiotapu and requires sampling there to isolate a representative.

Recent discoveries indicate that Cyanidiales are not as limited to acidic geothermal sites as once thought, having been discovered in acidic mesophilic conditions in a cave in the Atacama desert[64], a burning waste coal heap in the Czech Republic [160] and in neutral/alkaline geothermal in Turkey [159]. So there is the potential to isolate and describe other novel Cyanidiales by sampling more usual locations such as acidic stream/soils or geothermal locations that are more neutral or alkaline.

There is a breadth of possible future research pathways involving the strains isolated here and with Cyanidiales in New Zealand in general. I will be focused on generating the genomic data to determine the; urea metabolism pathway, presence of metal transporters for application in the proposed bioreactor and to resolve the taxonomy, to hopefully revise the nomenclature and amend the description. Collating this genomic data with the data gathered in Chapter 3 and 4, with retesting where required, and revising the data to a publishable standard in pursuit of publishing a scientific paper.

5.5 Conclusion

In conclusion, in this thesis Cyanidiales strains were sampled from three geothermal sites within the TVZ, NZ. The isolation method using solid Modified Allen's medium supplemented with glycerol and glucose allowed for detection of multiple Cyanidiales strains within single samples, due to observed differences in pigment and colony morphology. In total, 10 Cyanidiales representatives were isolated. The phenotypic characterization and phylogenetic analysis separated these 10 representatives into three lineages; one Eurasian *Galdieria maxima* lineage, two *Galdieria sulphuraria* lineages, one American and one Italian. The Italian *G. sulphuraria* lineage had not been previously observed in New Zealand[26,27]. This indicates that despite geographic isolation, the NZ Cyanidiales population is more diverse than previous sampling indicated[26] and other geothermal sites globally[26,32,78]. This is evidence for either multiple migration events to NZ and an undetermined long-distance dispersal mechanism, or that the diversity of other sites are underestimated due to under sampling or the methods of identification lacking resolution. The NZ *G. maxima* isolates were determined to be phylogenetically related to *G. maxima* IPPAS P507, but as obligate photoautotrophs with smaller oblong cells that are phenotypically distinct[36,61]. There is the possibility that NZ *G. maxima* represents a stepping stone between *Cyanidium* and *G. maxima,* which therefore presents the opportunity to determine the evolutionary pathway of *G. maxima* and revise the nomenclature. Evidence collected in this thesis corrected the identification of the previously isolated *Cyanidium caldarium* NZ4[27] as a *Galdieria* strain. This thesis contributes evidence that furthers our understanding of the dispersal, evolution and phylogeny of Cyanidiales both in New Zealand and globally. Additionally, this thesis contributes to the research and development of New Zealand biotechnology by providing isolated and characterised Cyanidiales. Cyanidiales as photosynthetic acido-thermophilic microorganisms are unique and are applicable in biotechnologies such as, dye production, effluent treatment, recovery of metal and sequestering carbon emissions from flue gas.

Chapter 6 Appendices and Supplementary material

6.1 Media Recipes

6.1.1 V4 medium recipe[23,135]

Table 6.1 V4 recipe for 1L of liquid media[23,135]

Compound	Quantity per litre
NH_4Cl	0.4 g
KH_2PO_4	0.05 g
$MgSO_4 \cdot 7H_2O$	0.02 g
$CaCl_2 \cdot 2H_2O$	0.01
FeEDTA solution (6.1.5)	3 mL
Trace element for methanotrophs (6.1.3)	3 mL
Trace element for methanogens (6.1.4)	1 mL
1mM Cerium sulfate solution	200 uL
1mM Lanthanum Chloride solution	200 uL
Yeast extract	0.01 g

Media is adjusted to pH 2.5 with H_2SO_4 then autoclaved. Details preparation of solid medium is detailed in section 6.1.6.

6.1.2 Modified Allen's medium

Table 6.2 Modified Allen's recipe for 1L of liquid media modified from *Minoda et al.* 2004

Compound	Quantity per litre
$(NH_4)_2SO_4$	2.62 g
KH_2PO_4	0.54 g
$MgSO_4 \cdot 7H_2O$	0.5 g
$CaCl_2 \cdot 2H_2O$	0.14g
FeEDTA solution (6.1.5)	6 mL

Trace element for methanotrophs (6.1.3)	6 mL
Trace element for methanogens (6.1.4)	2 mL
D-Glucose*	3 g
Glycerol*	2.38 mL / 3g

Media is adjusted to pH 2.5 with H_2SO_4 then autoclaved. * IS only added for Mall-Ht medium Details preparation of solid medium is detailed in section 6.1.6.

6.1.3 Trace elements for Methanotroph

Table 6.3 Recipe for trace element solution optimised for Methanotrophs[23,135]

Compound	Quantity per mL
$ZnSO_4 \cdot 7\ H_2O$	0.44 mg (= 0.1 mg Zn)
$CuSO_4 \cdot 5\ H_2O$	0.20 mg (= 0.05 mg Cu)
$MnCl \cdot 4H_2O$	0.19 mg
$Na_2MoO_4 \cdot 2\ H_2O$	0.06 mg (= 0.024 mg Mo)
H_3BO_3	0.10 mg (= 0.02 mg B)
$CoCl_2 \cdot 6\ H_2O$	0.08 mg (= .,02 mg Co)

6.1.4 Trace elements for Methanogens

Table 6.4 Recipe for trace element solution optimised for Methanogen modified from Wolins *et al.* 1963

Compound	Quantity per litre
Nitrilotriacetic acid	1.5 g
$Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$	0.2 g
Na_2SeO_3	0.2 g
$CoCl_2 \cdot 6H_2O$	0.1 g
$MnSO_4 \cdot 2H_2O$	0.1 g
$Na_2MoO_4 \cdot 2H_2O$	0.1 g
$Na_2WO_4 \cdot 2H_2O$	0.1 g

$ZnSO_4 \cdot 7H_2O$	0.1 g
$AlCl_3 \cdot 6H_2O$	0.04 g
$NiCl_2 \cdot 6H_2O$	0.025 g
H_3BO_3	0.01 g
$CuSO_4 \cdot 5H_2O$	0.01 g

To prepare the trace element solution, dissolve the nitrilotriacetic acid in 800 ml of water and adjust the pH to 6.5 with KOH. Then dissolve the minerals in order, adjust the pH to 7.0, and bring the volume to 1 liter.

6.1.5 Iron EDTA solution

Preparation of 1L of iron chelate solution

1) 1.54 g of $FeSO_4 \cdot 7\ H_2O$ was added to 400mL of MillQ water. Then was heated and stirred till salt completely dissolved.
2) 2.06g of Na_2EDTA was added to 400Ml of MillQ water. Then was heated and stirred till salt completely dissolved.
3) Solutions from 1) and 2) were combined and made up 1L. Final solution was wrapped in aluminium foil and stored at 4°C

6.1.6 Preparation of Solid media

For the solid medium to set correctly the gelling agent and nutrient solutions had to be autoclaved separately.
To do this nutrient solution was made up according to the recipes (see above) for 1 litre medium but were made up a final volume of 500mL. An additional 1L Schott bottle was filled 500mL of MillQ water and placed on a stirrer. The stirrer was adjusted so there was a large vortex then 15g of Phytagel was added steadily but rapidly to avoid clumps. Both solutions were autoclaved separately then ,aseptically, without allowing either solution to cool the nutrient solution was slowly poured (to avoid bubbles) into Schott bottle with the molten Phytagel. Plates were poured immediately after.

6.2 Chapter 4 Supplementary material

6.2.1 Primers tested

Table 6.4 All PCR primers tested

Name	**Sequence (5'-3')**

9F	AGAGTTTGATCMTGGCTCAG
1492R	GGHTACCTTGTTACGACTT
Calmo140F	GAA KCR GAR TTR MGA GAR ATG AT
Calmo160F	GAR ATG ATH GCV GAR GTW GA
Calmo510R	CTT CMT CBG TAA GYT TTT CTC C
Calmo580R	ATC YGC TTC RCG AAT CAT TTC
psaA130F	AACWACWACTTGGATTTGGAA
psaA180F	GATAGTCAWACHAGTTCWTTAGA
psaA1600R	GCATGAATATGRTGWACCAT
psaA1760R	CCTCTWCCWGGWCCATCRCAWGG
psbA-F	ATGACTGCTACTTTAGAAAGACG
psbA-R2	TCATGCATWACTTCCATACCTA
psbA-R1	GCTAAATCTARWGGGAAGTTGTG
RbcL1F	AACCTTTCATGCGTTGGAGAGA
RbcL1R	CCTGCATGAATACCACCAGAAGC
CdmF	GTCAGAGGTGAAATTCTTGGATTTA
CdmR	AAGGGCAGGGACGTAATCAACG

6.2.2 Accession numbers for phylogenetic analysis

Table 6.5 Accession numbers of all *rbc*L sequences used in phylogenetic analysis of *rbc*L

Organism	Strain	Accession
Bangia fuscopurpurea	SAG 59.81	AY119771.1
Bangiopsis subsimplex	PR21	AY119772.1
Cyanidiales sp. CCMEE 5506 EX	CCMEE 5506 EX	EF675146.1
Cyanidiales sp. CCMEE 5507 EX	CCMEE 5507 EX	EF675160.1
Cyanidiales sp. CCMEE 5508 DRG	CCMEE 5508 DRG	EF675135.1
Cyanidiales sp. CCMEE 5576 LEM	CCMEE 5576 LEM	EF675130.1
Cyanidiales sp. CCMEE 5584 NYM	CCMEE 5584 NYM	EF675164.1
Cyanidiales sp. CCMEE 5585 NYM	CCMEE 5585 NYM	EF675152.1
Cyanidiales sp. CCMEE 5593 OBS	CCMEE 5593 OBS	EF675124.1
Cyanidiales sp. CCMEE 5610 SCR	CCMEE 5610 SCR	EF675125.1
Cyanidiales sp. CCMEE 5625 IMP	CCMEE 5625 IMP	EF675158.1
Cyanidiales sp. CCMEE 5639 EXB	CCMEE 5639 EXB	EF675127.1
Cyanidiales sp. CCMEE 5640 EXB	CCMEE 5640 EXB	EF675137.1
Cyanidioschyzon merolae	10D	NC_004799.1
Cyanidioschyzon merolae	DBV 001 NAPS	AY119766.1
Cyanidioschyzon merolae	DBV 202 NAMN	AY541296.1
Cyanidium caldarium	DBV 019 SIPE	AY541297.1
Cyanidium caldarium	DBV 182 APAS	AY541299.1
Cyanidium caldarium	DBV 182 JAVA	AY541298.1
Cyanidium sp. MonteRotaro 19	MonteRotaro 19	AY541300.1

Cyanidium sp. MonteRotaro 20	MonteRotaro 20	AY541301.1
Galdieria daedala	IPPAS P508	AY541302.1
Galdieria maxima	IPPAS P507	AY391370.1
Galdieria partita	IPPAS P500	AB018008.1
Galdieria sp. CCMEE 5511	CCMEE 5511 EX	EF675174.1
Galdieria sp. CCMEE 5572	CCMEE 5572 EX	EF675182.1
Galdieria sp. CCMEE 5573	CCMEE 5573 EX	EF675171.1
Galdieria sp. CCMEE 5609	CCMEE 5609 SYL	EF675144.1
Galdieria sp. CCMEE 5620	CCMEE 5620 UTEX	EF675170.1
Galdieria sp. CCMEE 5631	CCMEE 5631 EX	EF675140.1
Galdieria sp. CCMEE 5657	CCMEE 5657 OWA	EF675139.1
Galdieria sp. CCMEE 5658	CCMEE 5658 OWA	EF675162.1
Galdieria sp. CCMEE 5659	CCMEE 5659 OWA	EF675138.1
Galdieria sp. CCMEE 5660	CCMEE 5660 NAK	EF675156.1
Galdieria sp. CCMEE 5661	CCMEE 5661 NAK	EF675150.1
Galdieria sp. CCMEE 5662	CCMEE 5662 NAK	EF675154.1
Galdieria sp. CCMEE 5663	CCMEE 5663 NAK	EF675159.1
Galdieria sp. CCMEE 5664	CCMEE 5664 KUS	EF675145.1
Galdieria sp. CCMEE 5665	CCMEE 5665 KUS	EF675129.1
Galdieria sp. CCMEE 5667	CCMEE 5667 KUS	EF675151.1
Galdieria sp. CCMEE 5668	CCMEE 5668 OWA	EF675136.1
Galdieria sp. CCMEE 5669	CCMEE 5669 OWA	EF675126.1
Galdieria sp. CCMEE 5670	CCMEE 5670 OWA	EF675148.1
Galdieria sp. CCMEE 5671	CCMEE 5671 OWA	EF675133.1
Galdieria sp. CCMEE 5672	CCMEE 5672 OWA	EF675131.1
Galdieria sp. CCMEE 5673	CCMEE 5673 OWA	EF675141.1
Galdieria sp. CCMEE 5674	CCMEE 5674 NAK	EF675153.1
Galdieria sp. CCMEE 5675	CCMEE 5675 NAK	EF675155.1
Galdieria sp. CCMEE 5676	CCMEE 5676 KUS	EF675143.1
Galdieria sp. CCMEE 5677	CCMEE 5677 KUS	EF675132.1
Galdieria sp. CCMEE 5678	CCMEE 5678 KUS	EF675168.1
Galdieria sp. CCMEE 5679	CCMEE 5679 KUS	EF675163.1
Galdieria sp. CCMEE 5680	CCMEE 5680 KUS	EF675167.1
Galdieria sp. CCMEE 5681	CCMEE 5681 KUS	EF675157.1
Galdieria sp. CCMEE 5702	CCMEE 5702 WM	EF675169.1
Galdieria sp. CCMEE 5703	CCMEE 5703 WM	EF675165.1
Galdieria sp. CCMEE 5704	CCMEE 5704 WM	EF675128.1
Galdieria sp. CCMEE 5705	CCMEE 5705 ROTO	EF675166.1
Galdieria sp. CCMEE 5706	CCMEE 5706 CR	EF675177.1
Galdieria sp. CCMEE 5707	CCMEE 5707 WTP	EF675181.1
Galdieria sp. CCMEE 5708	CCMEE 5708 WKA	EF675172.1
Galdieria sp. CCMEE 5709	CCMEE 5709 WKA	EF675161.1
Galdieria sp. CCMEE 5710	CCMEE 5710 WMC	EF675183.1

Galdieria sp. CCMEE 5711	CCMEE 5711 WI	EF675173.1
Galdieria sp. CCMEE 5712	CCMEE 5712 CR	EF675178.1
Galdieria sp. CCMEE 5713	CCMEE 5713 WTP	EF675147.1
Galdieria sp. CCMEE 5714	CCMEE 5714 WMC	EF675180.1
Galdieria sp. CCMEE 5715	CCMEE 5715 WKA	EF675149.1
Galdieria sp. CCMEE 5716	CCMEE 5716 CR	EF675142.1
Galdieria sp. CCMEE 5717	CCMEE 5717 RT	EF675176.1
Galdieria sp. CCMEE 5719	CCMEE 5719 WTP	EF675175.1
Galdieria sp. CCMEE 5720	CCMEE 5720 WI	EF675134.1
Galdieria sulphuraria	DBV 009 VTNE	AY119768.1
Galdieria sulphuraria	DBV 015 NAFG	AY541305.1
Galdieria sulphuraria	DBV 021 MEVU	AY541307.1
Galdieria sulphuraria	DBV 063 AGCS	AY119769.1
Galdieria sulphuraria	SAG 108.79	AY119767.1

Table 6.6 Accession numbers of all 18SrRNA gene sequences used 18S rRNA phylogenetic analysis

Name	Accession
Bangiopsis subsimplex	AF168627.1
Compsopogon caeruleus	AF342748.1
CCMEE 5506 EX	EF675087.1
CCMEE 5508 DRG	EF675085.1
CCMEE 5562 APP	EF675089.1
CCMEE 5564 CYA	EF675099.1
CCMEE 5576 LEM	EF675084.1
CCMEE 5580 MUD	EF675088.1
CCMEE 5584 NYM	EF675100.1
CCMEE 5585 NY	EF675079.1
CCMEE 5602 RC	EF675086.1
CCMEE 5605 RG	EF675076.1
CCMEE 5610 SCR	EF675083.1
CCMEE 5611 SCR	EF675074.1
CCMEE 5614 SOUR	EF675081.1
CCMEE 5617 STAR	EF675075.1
CCMEE 5624 IMP	EF675082.1
CCMEE 5625 HI	EF675078.1
CCMEE 5651 HRT	EF675077.1
CCMEE 5652 HRT	EF675080.1
CCMEE 5654 HLSA	EF675071.1
Cyanidioschyzon merolae strain DBV201	AF441376.1
Cyanidioschyzon merolae strain 10D	XR_002461616.1
Cyanidium caldarium_55B	AB091232.1
Cyanidium caldarium RK1(61D)	AB090833.1

Galdieria daedala_508	AF441362.1
Galdieria maxima_507	AF441367.1
Galdieria partita_500	AF441368.1
Galdieria partita_SAG107.79	AB091229.1
CCMEE 5511 EX	EF675094.1
CCMEE 5572 EX	EF675073.1
CCMEE 5573 EX	EF675092.1
CCMEE 5574 EX	EF675098.1
CCMEE 5620 UTEX	EF675090.1
CCMEE 5657 OWA	EF675091.1
CCMEE 5659 OWA	EF675097.1
CCMEE 5672 OWA	EF675095.1
CCMEE 5675 NAK	EF675093.1
CCMEE 5678 KUS	EF675096.1
CCMEE 5680 KUS	EF675072.1
CCMEE 5704 WM	EF675103.1
CCMEE 5706 CR	EF675106.1
CCMEE 5709 WKA	EF675101.1
CCMEE 5711 WI	EF675105.1
CCMEE 5716 CR	EF675104.1
CCMEE 5720 WI	EF675102.1
Galdieria sulphuraria_074W	AF441369.1
Galdieria sulphuraria_AZ	AF441360.1
Galdieria sulphuraria_Gs871-YNP	KP167586.1
Galdieria sulphuraria_MSh	AF441371.1
Galdieria sulphuraria SAG108.79	AF342747.1
Rhodochaete parvula	AF139462.1

Table 6.7 Accession numbers of all gene sequences used in the phylogenetic analysis of the three plastid protein gene concatenated dataset

Organism	**Strain**	**psaA**	**psbA**	**rbcL**
Bangia atropurpurea	SAG 33.94	AY119698.1	AY119734.1	AY119770.1
Bangia fuscopurpurea	SAG 59.81	AY119699.1	AY119735.1	AY119771.1
Bangiopsis subsimplex	PR21	AY119700.1	AY119736.1	AY119772.1
Compsopogon caeruleus	SAG 36.94	AY119701.1	AY119737.1	AF087116.1
Compsopogon caeruleus	SAG 36.94	AY119701.1	AY119737.1	AF087116.1
Cyanidioschyzon merolae	DBV 001 NAPS	AY119694.1	AY119730.1	AY119766.1
Cyanidioschyzon merolae	10D	NC_004799.1	NC_004799.1	NC_004799.1
Cyanidioschyzon merolae	DBV 201 JAVA	AY119693.1	AY119729.1	AY119765.1
Cyanidium caldarium	RK1	NC_001840.1	NC_001840.1	NC_001840.1
Cyanidium caldarium	DBV 019 SIPE	AY541281.1	AY541289.1	AY541297.1

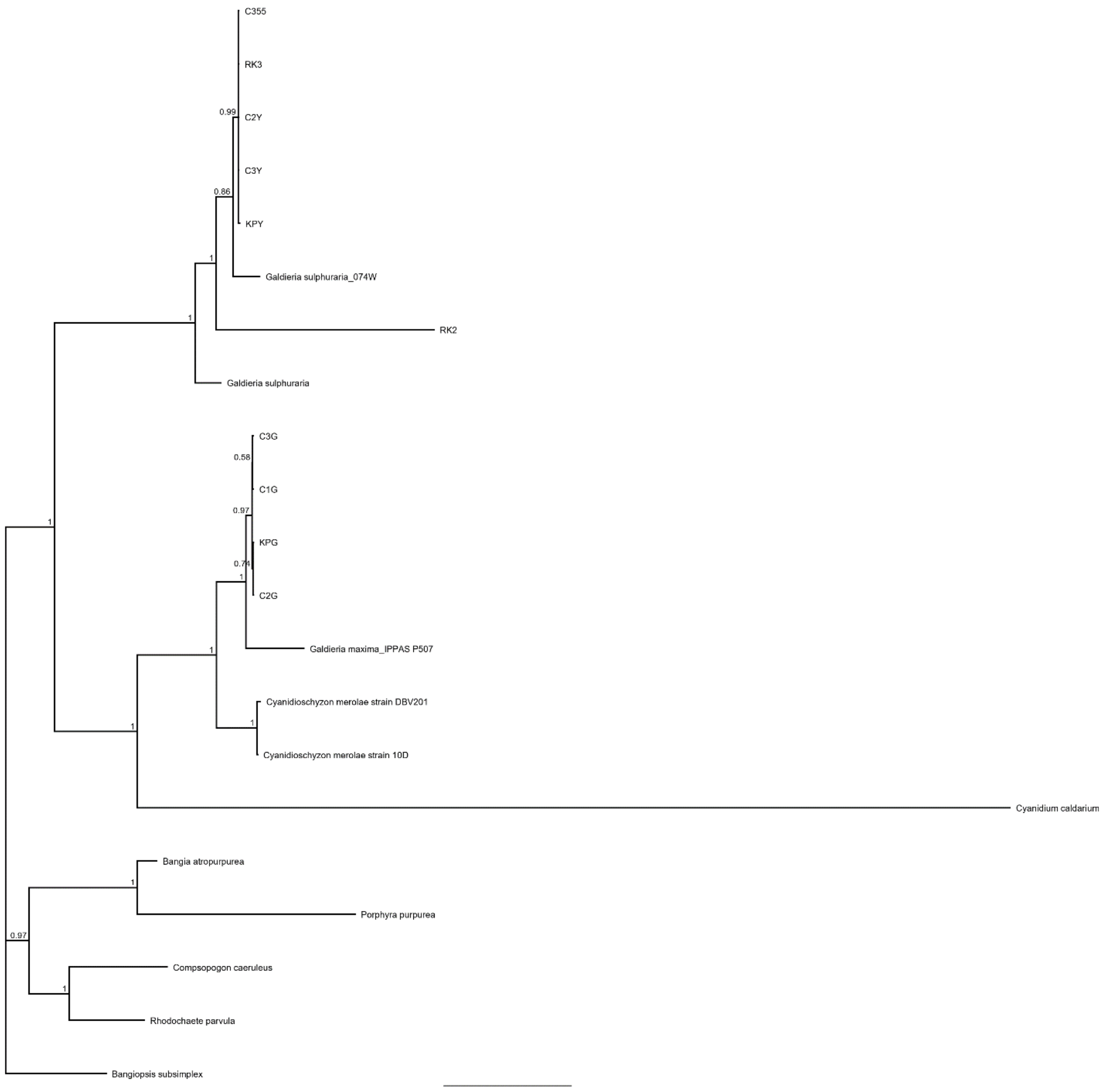
Cyanidium caldarium	DBV 182 JAVA	AY541282.1	AY541290.1	AY541298.1
Cyanidium sp. Monte Rotaro	Monte Rotaro	AY391362.1	AY391365.1	AY391368.1
Cyanidium sp. Sybil cave	Sybil cave	AY391363.1	AY391366.1	AY391369.1
Galdieria daedala	IPPAS P508	AY541283.1	AY541291.1	AY541302.1
Galdieria maxima	ACUF710	KX501186.1	KX501200.1	KX501173.1
Galdieria maxima	ACUF769	KX501191.1	KX501205.1	KX501179.1
Galdieria maxima	ACUF773	KX501192.1	KX501206.1	KX501180.1
Galdieria maxima	IPPAS P507	AY391364.1	AY391367.1	AY391370.1
Galdieria partita	IPPAS P500	AY541284.1	AY541292.1	AB018008.1
Galdieria phlegrea	ACUF738	KX501189.1	KX501203.1	KX501176.1
Galdieria phlegrea	ACUF780	KX501194.1	KX501208.1	KX501182.1
Galdieria sulphuraria	074W	NC_024665.1	NC_024665.1	NC_024665.1
Galdieria sulphuraria	ACUF388	KC883922.1	KC883967.1	KC883815.1
Galdieria sulphuraria	ACUF461	KC883919.1	KC883961.1	KC883873.1
Galdieria sulphuraria	DBV 011 CEMD	AY541286.1	AY541293.1	AY541303.1
Galdieria sulphuraria	DBV 012 BNTE	AY541288.1	AY541295.1	AY541310.1
Galdieria sulphuraria	DBV 018 NASC	AY541287.1	AY541294.1	AY541304.1
Galdieria sulphuraria	DBV 063 AGCS	AY119697.1	AY119733.1	AY119769.1
Galdieria sulphuraria	SAG 108.79	AY119695.1	AY119731.1	AY119767.1
Porphyra purpurea	Avonport	NC_000925.1	NC_000925.1	NC_000925.1
Rhodochaete parvula	UTEX LB 2715	NC_031180.2	AY119743.1	AY119777.1

Table 6.8 Accession numbers of all gene sequences used in the phylogenetic analysis of the Five-gene concatenated dataset

Organism	**Strain**	**18S rRNA**	**16S rRNA plastid**	**psaA**	**psbA**	**rbcL**
Bangia atropurpurea	SAG 33.94	D88387.1	AF545616.1	AY119698.1	AY119734.1	AY119770.1
Bangiopsis subsimplex	PR21	AF168627.1	AF545620.1	AY119700.1	AY119736.1	AY119772.1
Compsopogon caeruleus	SAG:36.94	AF342748.1	AF170713.1	AY119701.1	AY119737.1	AF087116.1
Cyanidioschyzon merolae	10D	XR_002461616.1	NC_004799.1	NC_004799.1	NC_004799.1	NC_004799.1
Cyanidioschyzon merolae	DBV201	AF441376.1	AF545617.1	AY119693.1	AY119729.1	AY119765.1
Cyanidium caldarium	RK1	AB090833.1	NC_001840.1	NC_001840.1	NC_001840.1	NC_001840.1
Galdieria maxima	P507	AF441367.1	AY391361.1	AY391364.1	AY391367.1	AY391370.1

Galdieria sulphuraria	SAG108.79	AF342747.1	AF170718.1	AY119695.1	AY119731.1	AY119767.1
Galdieria sulphuraria	074W	AF441369.1	NC_024665.1	NC_024665.1	NC_024665.1	NC_024665.1
Porphyra purpurea	Avonport	AF362362.1	NC_000925.1	NC_000925.1	NC_000925.1	NC_000925.1
Rhodochaete parvula	UTEX LB 2715	AF139462.1	AF545623.1	NC_031180.2	AY119743.1	AY119777.1

6.3.2 Supplementary trees



Supplementary figure 6.1: The phylogeny of New Zealand Cyanidiales inferred from Bayesian analysis of the combined DNA sequences of *psa*A, *psb*A, *rbc*L, 18S rRNA and plastid 16SrRNA genes with nucleotide from 3rd codon position protein coding sequences excluded. Bayesian posterior probabilities are shown on above the branches as decimals. Sequences from this thesis are indicated with red brackets and described lineages are labelled and indicated with black brackets.

References

1. Etheridge, D. M. *et al.* Natural and anthropogenic changes in atmospheric CO2 over the last 1000 years from air in Antarctic ice and firn. *J. Geophys. Res. D: Atmos.* **101**, 4115–4128 (1996).
2. Dolman, H., Valentini, R. & Freibauer, A. *The Continental-Scale Greenhouse Gas Balance of Europe*. (Springer Science & Business Media, 2008).
3. Raupach, M. R. *et al.* Global and regional drivers of accelerating CO2 emissions. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10288–10293 (2007).
4. New Zealand Greenhouse Gas Inventory 2019.
5. Zeebe, R. E., Zachos, J. C., Caldeira, K. & Tyrrell, T. Oceans. Carbon emissions and acidification. *Science* **321**, 51–52 (2008).
6. Matthews, H. D., Gillett, N. P., Stott, P. A. & Zickfeld, K. The proportionality of global warming to cumulative carbon emissions. *Nature* **459**, 829–832 (2009).
7. Jacobson, M. Z. Review of solutions to global warming, air pollution, and energy security. *Energy Environ. Sci.* **2**, 148–173 (2009).
8. Li, Y., Horsman, M., Wu, N., Lan, C. Q. & Dubois-Calero, N. Biofuels from microalgae. *Biotechnol. Prog.* **24**, 815–820 (2008).
9. Khan, M. I., Shin, J. H. & Kim, J. D. The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microb. Cell Fact.* **17**, 36 (2018).
10. le B. Williams, P. J. & Laurens, L. M. L. Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy Environ. Sci.* **3**, 554–590 (2010).
11. Grassini, P., Eskridge, K. M. & Cassman, K. G. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun.* **4**, 2918 (2013).
12. Graziani, G. *et al.* Microalgae as human food: chemical and nutritional characteristics of the thermo-acidophilic microalga Galdieria sulphuraria. *Food Funct.* **4**, 144–152 (2013).
13. Mata, T. M., Martins, A. A. & Caetano, N. S. Microalgae for biodiesel production and other applications: A review. *Renewable Sustainable Energy Rev.* **14**, 217–232 (2010).
14. Varshney, P., Mikulic, P., Vonshak, A., Beardall, J. & Wangikar, P. P. Extremophilic micro-algae and their potential contribution in biotechnology. *Bioresour. Technol.* **184**, 363–372 (2015).
15. Spolaore, P., Joannis-Cassan, C., Duran, E. & Isambert, A. Commercial applications of microalgae. *J. Biosci. Bioeng.* **101**, 87–96 (2006).
16. Pulz, O. & Gross, W. Valuable products from biotechnology of microalgae. *Appl. Microbiol. Biotechnol.* **65**, 635–648 (2004).
17. Becker, W. Microalgae in Human and Animal Nutrition. *Handbook of Microalgal Culture* 312–351 doi:10.1002/9780470995280.ch18.
18. Stocker, T. F. *et al.* Climate change 2013: The physical science basis. (2013).
19. Gregory, J. *et al.* Climate change 2007: the physical science basis. (2007).
20. Carere, C. R. *MBIE Endeavour Smart Idea Grant. Flipping the Paradigm: feeding methane to cows*. (2016).
21. Tchinda, D. *et al.* Single-step treatment of primary effluent by Galdieria sulphuraria:

Removal of biochemical oxygen demand, nutrients, and pathogens. *Algal Research* **42**, (2019).
22. Op den Camp, H. J. M. *et al.* Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia. *Environ. Microbiol. Rep.* **1**, 293–306 (2009).
23. Dunfield, P. F. *et al.* Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature* **450**, 879–882 (2007).
24. Brock, T. D. & Brock, M. L. THE ALGAE OF WAIMANGU CAULDRON (NEW ZEALAND): DISTRIBUTION IN RELATION TO pH 1. *J. Phycol.* **6**, 371–375 (1970).
25. Brock, T. D. & Brock, M. L. Microbiological studies of thermal habitats of the central volcanic region, North Island, New Zealand. *N. Z. J. Mar. Freshwater Res.* **5**, 233–258 (1971).
26. Toplin, J. A., Norris, T. B., Lehr, C. R., McDermott, T. R. & Castenholz, R. W. Biogeographic and phylogenetic diversity of thermoacidophilic cyanidiales in Yellowstone National Park, Japan, and New Zealand. *Appl. Environ. Microbiol.* **74**, 2822–2833 (2008).
27. Donachie, S. P., Christenson, B. W., Kunkel, D. D., Malahoff, A. & Alam, M. Microbial community in acidic hydrothermal waters of volcanically active White Island, New Zealand. *Extremophiles* **6**, 419–425 (2002).
28. Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G. & Bhattacharya, D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809–818 (2004).
29. Yoon, H. S., Muller, K. M., Sheath, R. G., Ott, F. D. & Bhattacharya, D. DEFINING THE MAJOR LINEAGES OF RED ALGAE (RHODOPHYTA)1. *J. Phycol.* **42**, 482–492 (2006).
30. Merola, A. *et al.* Revision of Cyanidium caldarium. Three species of acidophilic algae. *Plant Biosyst.* **115**, 189–195 (1981).
31. Albertano, P., Ciniglia, C., Pinto, G. & Pollio, A. The taxonomic position of Cyanidium, Cyanidioschyzon and Galdieria: an update. *Hydrobiologia* **433**, 137–143 (2000).
32. Ciniglia, C., Yang, E. C., Pollio, A., Pinto, G. & Iovinella, M. Cyanidiophyceae in Iceland: plastid rbcL gene elucidates origin and dispersal of extremophilic Galdieria sulphuraria and G. maxima (Galdieriaceae, Rhodophyta). (2014).
33. Tansey, M. R. & Brock, T. D. The upper temperature limit for eukaryotic organisms. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 2426–2428 (1972).
34. Rothschild, L. J. & Mancinelli, R. L. Life in extreme environments. *Nature* **409**, 1092–1101 (2001).
35. Clarke, A. The thermal limits to life on Earth. *Int. J. Astrobiology* **13**, 141–154 (2014).
36. Pinto, G. *et al.* Comparative approaches to the taxonomy of the genus Galdieria merola (Cyanidiales, Rhodophyta). *Cryptogamie-Algologie* **24**, 13–32 (2003).
37. Seckbach, J., Baker, F. A. & Shugarman, P. M. Algae thrive under pure CO2. *Nature* **227**, 744–745 (1970).
38. Ciniglia, C., Yoon, H. S., Pollio, A., Pinto, G. & Bhattacharya, D. Hidden biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.* **13**, 1827–1838 (2004).
39. Yoon, H. S., Hackett, J. D. & Bhattacharya, D. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11724–11729 (2002).
40. Kuroiwa, T. The primitive red algae Cyanidium caldarium and Cyanidioschyzon merolae as model system for investigating the dividing apparatus of mitochondria and plastids.

Bioessays **20**, 344–354 (1998).
41. Misumi, O. *et al.* Cyanidioschyzon merolae genome. A tool for facilitating comparable studies on organelle biogenesis in photosynthetic eukaryotes. *Plant Physiol.* **137**, 567–585 (2005).
42. Matsuzaki, M. *et al.* Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. *Nature* **428**, 653–657 (2004).
43. Nozaki, H. *et al.* A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga Cyanidioschyzon merolae. *BMC Biol.* **5**, 28 (2007).
44. Schönknecht, G. *et al.* Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* **339**, 1207–1210 (2013).
45. Qiu, H. *et al.* Adaptation through horizontal gene transfer in the cryptoendolithic red alga Galdieria phlegrea. *Curr. Biol.* **23**, R865–6 (2013).
46. Rossoni, A. W. *et al.* The genomes of polyextremophilic cyanidiales contain 1% horizontally transferred genes with diverse adaptive functions. *Elife* **8**, (2019).
47. Sydney, E. B. *et al.* Biomolecules from extremophile microalgae: From genetics to bioprocessing of a new candidate for large-scale production. *Process Biochem.* (2019) doi:10.1016/j.procbio.2019.09.012.
48. Allen, M. B. Studies with Cyanidium caldarium, an anomalously pigmented chlorophyte. *Arch. Mikrobiol.* **32**, 270–277 (1959).
49. De Luca, P. & Taddei, R. On the necessity of a systematic revision of the thermal acidophilic alga «Cyanidium caldarium»(Tilden) Geitler. *Webbia* (1976).
50. Ott, F. D. & Seckbach, J. A review on the taxonomic position of the algal genus Cyanidium Geitler 1933 and its ecological cohorts Galdieria Merola in Merola *et al.* 1981 and Cyanidioschyzon De Luca, Taddei and Varano 1978. in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells* (ed. Seckbach, J.) 113–132 (Springer Netherlands, 1994).
51. Bailey, R. W. & Staehelin, L. A. The chemical composition of isolated cell walls of Cyanidium caldarium. *J. Gen. Microbiol.* **54**, 269–276 (1968).
52. Ott, F. D. & Seckbach, J. New classification for the genus Cyanidium Geitler 1933. in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells* (ed. Seckbach, J.) 145–152 (Springer Netherlands, 1994).
53. Seckbach, J. The natural history of Cyanidium (Geitler 1933): past and present perspectives. in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells* (ed. Seckbach, J.) 99–112 (Springer Netherlands, 1994).
54. Boenzi, D., Luca, P. D. & Taddei, R. Fatty acids in \guillemotleftCyanidium\guillemotright. *Plant Biosyst.* **111**, 129–134 (1977).
55. De Luca, P., Taddei, R. & Varano, L. \guillemotleftCyanidioschyzon merolae\guillemotright: a new alga of thermal acidic environments. *Webbia* **33**, 37–44 (1978).
56. Kuroiwa, T. *et al.* Comparison of ultrastructures between the ultra-small eukaryote Cyanidioschyzon merolae and Cyanidium caldarium. *Cytologia* **59**, 149–158 (1994).
57. Barbier, G. *et al.* Comparative genomics of two closely related unicellular thermo-acidophilic red algae, Galdieria sulphuraria and Cyanidioschyzon merolae, reveals the molecular basis of the metabolic flexibility of Galdieria sulphuraria and significant differences in carbohydrate metabolism of both algae. *Plant Physiol.* **137**, 460–474 (2005).

58. Seckbach, J. SYSTEMATIC PROBLEMS WITH CYANIDIUM CALDARIUM AND GALDIERIA SULPHURARIA AND THEIR IMPLICATIONS FOR MOLECULAR BIOLOGY STUDIES1. *J. Phycol.* **27**, 794–796 (1991).
59. Gross, W. Revision of Comparative Traits for the Acido- and Thermophilic Red Algae Cyanidium and Galdieria. in *Enigmatic Microorganisms and Life in Extreme Environments* (ed. Seckbach, J.) 437–446 (Springer Netherlands, 1999).
60. Seckbach, J. The first eukaryotic cells — Acid hot-spring algae. *J. Biol. Phys.* **20**, 335–345 (1995).
61. Sentsova, O. Y. The study of Cyanidiophyceae in Russia. in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells* (ed. Seckbach, J.) 167–174 (Springer Netherlands, 1994).
62. Hoffmann, L. & Mayele, B. Ultrastructure of the plasmalemma of the cave Cyanidium (Rhodophyta). *Phycologia* **32**, 307–309 (1993).
63. Hoffmann, L. Cyanidium-like algae from caves. in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells* (ed. Seckbach, J.) 175–182 (Springer Netherlands, 1994).
64. Azúa-Bustos, A. *et al.* Ancient photosynthetic eukaryote biofilms in an Atacama Desert coastal cave. *Microb. Ecol.* **58**, 485–496 (2009).
65. Cennamo, P., Marzano, C., Ciniglia, C. & Pinto, G. A survey of the algal flora of anthropogenic caves of Campi Flegrei (Naples, Italy) archeological district. *J. Cave Karst Stud.* (2012).
66. Darienko, T. & Hoffmann, L. Subaerial algae and cyanobacteria from the archaeological remains of Carthage (Tunisia), including the record of a species of Cyanidium (Rhodophyta). *Arch. Hydrobiol. Suppl. Algol. Stud.* **135**, 41–60 (2010).
67. Ciniglia, C., Pinto, G. & Pollio, A. Cyanidium from caves: a reinstatement of Cyanidium chilense Schwabe (Cyanidiophytina, Rhodophyta). *Phytotaxa* **295**, 86–88 (2017).
68. Cozzolino, S., Caputo, P., De Castro, O., Moretti, A. & Pinto, G. Molecular variation in Galdieria sulphuraria (Galdieri) Merola and its bearing on taxonomy. *Hydrobiologia* **433**, 145–151 (2000).
69. Vítová, M., Goecke, F., Sigler, K. & Řezanka, T. Lipidomic analysis of the extremophilic red alga Galdieria sulphuraria in response to changes in pH. *Algal Research* **13**, 218–226 (2016).
70. Sakurai, T. *et al.* Profiling of lipid and glycogen accumulations under different growth conditions in the sulfothermophilic red alga Galdieria sulphuraria. *Bioresour. Technol.* **200**, 861–866 (2016).
71. López, G. *et al.* Production of Polyunsaturated Fatty Acids and Lipids from Autotrophic, Mixotrophic and Heterotrophic cultivation of Galdieria sp. strain USBA-GBX-832. *Sci. Rep.* **9**, 10791 (2019).
72. Gross, W., Heilmann, I., Lenze, D. & Schnarrenberger, C. Biogeography of the Cyanidiaceae (Rhodophyta) based on 18S ribosomal RNA sequence data. *Eur. J. Phycol.* **36**, 275–280 (2001).
73. Yoon, H. S., Hackett, J. D., Pinto, G. & Bhattacharya, D. The single, ancient origin of chromist plastids. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15507–15512 (2002).
74. Yang, E. C. *et al.* Divergence time estimates and the evolution of major lineages in the

florideophyte red algae. *Sci. Rep.* **6**, 21361 (2016).
75. Yoon, H. S. *et al.* Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). *BMC Evol. Biol.* **6**, 78 (2006).
76. Qiu, H., Price, D. C., Yang, E. C., Yoon, H. S. & Bhattacharya, D. Evidence of ancient genome reduction in red algae (Rhodophyta). *J. Phycol.* **51**, 624–636 (2015).
77. Pinto, G., Ciniglia, C., Cascone, C. & Pollio, A. Species Composition of Cyanidiales Assemblages in Pisciarelli (Campi Flegrei, Italy) and Description of Galdieria Phlegrea SP. NOV. in *Algae and Cyanobacteria in Extreme Environments* (ed. Seckbach, J.) 487–502 (Springer Netherlands, 2007).
78. Hsieh, C.-J., Zhan, S. H., Lin, Y., Tang, S.-L. & Liu, S.-L. Analysis of rbcL sequences reveals the global biodiversity, community structure, and biogeographical pattern of thermoacidophilic red algae (Cyanidiales). *J. Phycol.* **51**, 682–694 (2015).
79. Barcytė, D. *et al.* Burning coal spoil heaps as a new habitat for the extremophilic red alga Galdieria sulphuraria. *Fottea* **18**, 19–29 (2018).
80. Castenholz, R. W. The effect of sulfide on the bluegreen algae of hot springs. I. New zealand and iceland 1. *J. Phycol.* **12**, 54–68 (1976).
81. Vogl, C. *et al.* Probabilistic analysis indicates discordant gene trees in chloroplast evolution. *J. Mol. Evol.* **56**, 330–340 (2003).
82. Glöckner, G., Rosenthal, A. & Valentin, K. The structure and gene repertoire of an ancient red algal plastid genome. *J. Mol. Evol.* **51**, 382–390 (2000).
83. Walker, J. J., Spear, J. R. & Pace, N. R. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* **434**, 1011–1014 (2005).
84. Lehr, C. R. *et al.* CYANIDIA (CYANIDIALES) POPULATION DIVERSITY AND DYNAMICS IN AN ACID-SULFATE-CHLORIDE SPRING IN YELLOWSTONE NATIONAL PARK. *J. Phycol.* **43**, 3–14 (2007).
85. Skorupa, D. J., Reeb, V., Castenholz, R. W., Bhattacharya, D. & McDermott, T. R. Cyanidiales diversity in Yellowstone National Park. *Lett. Appl. Microbiol.* **57**, 459–466 (2013).
86. Castenholz, R. W. & McDermott, T. R. The Cyanidiales: Ecology, Biodiversity, and Biogeography. in *Red Algae in the Genomic Age* (eds. Seckbach, J. & Chapman, D. J.) 357–371 (Springer Netherlands, 2010).
87. Brock, T. D. *Thermophilic Microorganisms and Life at High Temperatures*. (Springer Science & Business Media, 2012).
88. Ohta, N., Sato, N. & Kuroiwa, T. Structure and organization of the mitochondrial genome of the unicellular red alga Cyanidioschyzon merolae deduced from the complete nucleotide sequence. *Nucleic Acids Res.* **26**, 5190–5198 (1998).
89. Ohta, N. *et al.* Complete sequence and analysis of the plastid genome of the unicellular red alga Cyanidioschyzon merolae. *DNA Res.* **10**, 67–77 (2003).
90. Wendland, J. & Walther, A. Genome evolution in the eremothecium clade of the Saccharomyces complex revealed by comparative genomics. *G3* **1**, 539–548 (2011).
91. Gross, W. & Schnarrenberger, C. Heterotrophic Growth of Two Strains of the Acido-Thermophilic Red Alga Galdieria sulphuraria. *Plant Cell Physiol.* **36**, 633–638 (1995).
92. Weber, A. P. M. *et al.* EST-analysis of the thermo-acidophilic red microalga Galdieria sulphuraria reveals potential for lipid A biosynthesis and unveils the pathway of carbon

export from rhodoplasts. *Plant Mol. Biol.* **55**, 17–32 (2004).
93. Reed, R. H. Osmoacclimation in Bangia atropurpurea (Rhodophyta, Bangiales): the osmotic role of floridoside. *British Phycological Journal* **20**, 211–218 (1985).
94. Reed, R. H. Taxonomic implications of osmoacclimation in Cyanidium caldarium (Tilden) Geitler. *Phycologia* **22**, 351–354 (1983).
95. Li, S.-Y., Shabtai, Y. & Arad, S. FLORIDOSIDE AS A CARBON PRECURSOR FOR THE SYNTHESIS OF CELL-WALL POLYSACCHARIDE IN THE RED MICROALGA PORPHYRIDIUM SP.(RHODOPHYTA) 1. *J. Phycol.* **38**, 931–938 (2002).
96. Luca, P. & Moretti, A. FLORIDOSIDES IN CYANIDIUM CALDARIUM, CYANIDIOSCHYZON MEROLAE AND GALDIERIA SULPHURARIA(RHODOPHYTA, CYANIDIOPHYCEAE)1. *J. Phycol.* **19**, 368–369 (1983).
97. Nagashima, H. & Fukuda, I. Low molecular weight carbohydrates in Cyanidium caldarium and some related algae. *Phytochemistry* **20**, 439–442 (1981).
98. Nagashima, H. & Fukuda, I. Floridosides in unicellular hot spring algae. *Phytochemistry* **22**, 1949–1951 (1983).
99. Qiu, H., Yoon, H. S. & Bhattacharya, D. Red Algal Phylogenomics Provides a Robust Framework for Inferring Evolution of Key Metabolic Pathways. *PLoS Curr.* **8**, (2016).
100. Collén, J. *et al.* Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 5247–5252 (2013).
101. Bhattacharya, D. *et al.* Genome of the red alga Porphyridium purpureum. *Nat. Commun.* **4**, 1941 (2013).
102. Martin, W. F. Too Much Eukaryote LGT. *Bioessays* **39**, (2017).
103. Martin, W. F. Eukaryote lateral gene transfer is Lamarckian. *Nature ecology & evolution* vol. 2 754 (2018).
104. Danchin, E. G. J. Lateral gene transfer in eukaryotes: tip of the iceberg or of the ice cube? *BMC biology* vol. 14 101 (2016).
105. Chen, F., Zhang, Y. & Guo, S. Growth and phycocyanin formation of Spirulina platensis in photoheterotrophic culture. *Biotechnol. Lett.* **18**, 603–608 (1996).
106. Kuddus, M., Singh, P., Thomas, G. & Al-Hazimi, A. Recent developments in production and biotechnological applications of C-phycocyanin. *Biomed Res. Int.* **2013**, 742859 (2013).
107. Hu, J., Nagarajan, D., Zhang, Q., Chang, J.-S. & Lee, D.-J. Heterotrophic cultivation of microalgae for pigment production: A review. *Biotechnol. Adv.* **36**, 54–67 (2018).
108. Chen, F. & Zhang, Y. High cell density mixotrophic culture of Spirulina platensis on glucose for phycocyanin production using a fed-batch system. *Enzyme Microb. Technol.* **20**, 221–224 (1997).
109. Sloth, J. K., Wiebe, M. G. & Eriksen, N. T. Accumulation of phycocyanin in heterotrophic and mixotrophic cultures of the acidophilic red alga Galdieria sulphuraria. *Enzyme Microb. Technol.* **38**, 168–175 (2006).
110. Oesterhelt, C., Schmälzlin, E., Schmitt, J. M. & Lokstein, H. Regulation of photosynthesis in the unicellular acidophilic red alga Galdieria sulphuraria. *Plant J.* **51**, 500–511 (2007).
111. Moon, M. *et al.* Isolation and characterization of thermostable phycocyanin from Galdieria sulphuraria. *Korean J. Chem. Eng.* **31**, 490–495 (2014).
112. Rahman, D. Y., Sarian, F. D., van Wijk, A., Martinez-Garcia, M. & van der Maarel, M. J. E.

C. Thermostable phycocyanin from the red microalga Cyanidioschyzon merolae, a new natural blue food colorant. *J. Appl. Phycol.* **29**, 1233–1239 (2017).
113. Graverholt, O. S. & Eriksen, N. T. Heterotrophic high-cell-density fed-batch and continuous-flow cultures of Galdieria sulphuraria and production of phycocyanin. *Appl. Microbiol. Biotechnol.* **77**, 69–75 (2007).
114. Schmidt, R. A., Wiebe, M. G. & Eriksen, N. T. Heterotrophic high cell-density fed-batch cultures of the phycocyanin-producing red alga Galdieria sulphuraria. *Biotechnol. Bioeng.* **90**, 77–84 (2005).
115. Sloth, J. K., Jensen, H. C., Pleissner, D. & Eriksen, N. T. Growth and phycocyanin synthesis in the heterotrophic microalga Galdieria sulphuraria on substrates made of food waste from restaurants and bakeries. *Bioresour. Technol.* **238**, 296–305 (2017).
116. Ahlf, W. Recovery of metals from acid waste water by Cyanidium caldarium. *Appl. Microbiol. Biotechnol.* **28**, 512–513 (1988).
117. Minoda, A. *et al.* Recovery of rare earth elements from the sulfothermophilic red alga Galdieria sulphuraria using aqueous acid. *Appl. Microbiol. Biotechnol.* **99**, 1513–1519 (2015).
118. Ju, X. *et al.* Effective and selective recovery of gold and palladium ions from metal wastewater using a sulfothermophilic red alga, Galdieria sulphuraria. *Bioresour. Technol.* **211**, 759–764 (2016).
119. Barragan, M. H. & Harcum, S. W. The feasibility of using Cyanidium Caldarium to bioremediate copper from acid mine drainage. *Radioact. Waste Manage. Environ. Restor.* **22**, 205–222 (2000).
120. Fukuda, S.-Y., Yamamoto, R., Iwamoto, K. & Minoda, A. Cellular accumulation of cesium in the unicellular red alga Galdieria sulphuraria under mixotrophic conditions. *J. Appl. Phycol.* **30**, 3057–3061 (2018).
121. Selvaratnam, T. *et al.* Evaluation of a thermo-tolerant acidophilic alga, Galdieria sulphuraria, for nutrient removal from urban wastewaters. *Bioresour. Technol.* **156**, 395–399 (2014).
122. Henkanatte-Gedera, S. M. *et al.* Algal-based, single-step treatment of urban wastewaters. *Bioresour. Technol.* **189**, 273–278 (2015).
123. Henkanatte-Gedera, S. M. *et al.* Removal of dissolved organic carbon and nutrients from urban wastewaters by Galdieria sulphuraria: Laboratory to field scale demonstration. *Algal Research* **24**, 450–456 (2017).
124. Cheng, F. *et al.* Bio-crude oil from hydrothermal liquefaction of wastewater microalgae in a pilot-scale continuous flow reactor. *Bioresour. Technol.* **294**, 122184 (2019).
125. Cheng, F. *et al.* Hydrothermal liquefaction of high- and low-lipid algae: Bio-crude oil chemistry. *Appl. Energy* **206**, 278–292 (2017).
126. Pires, J. C. M., Alvim-Ferraz, M. C. M., Martins, F. G. & Simões, M. Carbon dioxide capture from flue gases using microalgae: Engineering aspects and biorefinery concept. *Renewable Sustainable Energy Rev.* **16**, 3043–3053 (2012).
127. Van Den Hende, S., Vervaeren, H. & Boon, N. Flue gas compounds and microalgae: (bio-)chemical interactions leading to biotechnological opportunities. *Biotechnol. Adv.* **30**, 1405–1424 (2012).
128. Lee, J.-S. *et al.* Effects of SO2 and NO on growth of Chlorella sp. KR-1. *Bioresour.*

Technol. **82**, 1–4 (2002).
129. Woodward, C. A., Macinnis, J. M., Lewis, S. N. & Greenbaum, E. Chemical interaction of flue gas components with the growth ofcyanidium caldarium. *Appl. Biochem. Biotechnol.* **34**, 819 (1992).
130. Hauck, J. T., Scierka, S. J. & Perry, M. B. Effects of simulated flue gas on growth of microalgae. *Preprints of Papers, American Chemical Society, Division of Fuel Chemistry* **41**, (1996).
131. Kurano, N. *et al.* Fixation and utilization of carbon dioxide by microalgal photosynthesis. *Energy Convers. Manage.* **36**, 689–692 (1995).
132. Rooke, J. C. *et al.* Prolonging the lifetime and activity of silica immobilised Cyanidium caldarium. *J. Colloid Interface Sci.* **356**, 159–164 (2011).
133. Ciniglia, C., Yoon, H. S., Pollio, A., Pinto, G. & Bhattacharya, D. Hidden biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.* **13**, 1827–1838 (2004).
134. Jones, B. & Renaut, R. W. GROWTH OF SILICEOUS SPICULES IN ACIDIC HOT SPRINGS, WAIOTAPU GEOTHERMAL AREA, NORTH ISLAND, NEW ZEALAND. *Palaios* **21**, 406–423 (2006).
135. Pol, A. *et al.* Rare earth metals are essential for methanotrophic life in volcanic mudpots. *Environ. Microbiol.* **16**, 255–264 (2014).
136. Minoda, A., Sakagami, R., Yagisawa, F., Kuroiwa, T. & Tanaka, K. Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, Cyanidioschyzon merolae 10D. *Plant Cell Physiol.* **45**, 667–671 (2004).
137. Svetashev, V. I., Vysotskii, M. V., Ivanova, E. P. & Mikhailov, V. V. Cellular Fatty Acids of Alteromonas Species. *Syst. Appl. Microbiol.* **18**, 37–43 (1995).
138. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
139. Vaskovsky, V. E., Kostetsky, E. Y. & Vasendin, I. M. A universal reagent for phospholipid analysis. *J. Chromatogr.* **114**, 129–141 (1975).
140. Miner, C. S. & Dalton, N. N. Chemical properties and Derivatives of Glycerol. (1953).
141. Fuggi, A., Di Martino Rigano, V., Vona, V. & Rigano, C. Nitrate and ammonium assimilation in algal cell-suspensions and related pH variations in the external medium, monitored by electrodes. *Plant Sci. Lett.* **23**, 129–138 (1981).
142. Iovinella, M., Carbone, D. A., Cioppa, D. & Davis, S. J. Prevalent pH Controls the Capacity of Galdieria Maxima to Use Ammonia and Nitrate as a Nitrogen Source. *bioRxiv* (2019).
143. Lowell, C. & Castenholz, R. W. The lowering of external pH in confined environments by thermo-acidophilic algae (class: Cyanidiophyceae). *Environ. Microbiol. Rep.* **5**, 660–664 (2013).
144. Wang, J., Rosov, T., Wensel, P., McGowen, J. & Curtis, W. R. A preliminary implementation of metabolic-based pH control to reduce CO2 usage in outdoor flat-panel photobioreactor cultivation of Nannochloropsis oceanica microalgae. *Algal Research* **18**, 288–295 (2016).
145. Wang, J. & Curtis, W. R. Proton stoichiometric imbalance during algae photosynthetic growth on various nitrogen sources: toward metabolic pH control. *J. Appl. Phycol.* **28**, 43–52 (2016).
146. Raven, J. A. BIOCHEMICAL DISPOSAL OF EXCESS H+ IN GROWING PLANTS? *New*

Phytol. **104**, 175–206 (1986).
147. Miflin, B. J. & Lea, P. J. The pathway of nitrogen assimilation in plants. *Phytochemistry* **15**, 873–885 (1976).
148. Tischner, R. Nitrate uptake and reduction in higher and lower plants. *Plant Cell Environ.* **23**, 1005–1024 (2000).
149. Novacky, A. & Fischer, E. Membrane potential changes during transport of glycine as a neutral amino acid and nitrate in Lemna gibba G 1. *FEBS Lett.* **88**, 264–267 (1978).
150. Sato, N. & Moriyama, T. Genomic and biochemical analysis of lipid biosynthesis in the unicellular rhodophyte Cyanidioschyzon merolae: lack of a plastidic desaturation pathway results in the …. *Eukaryot. Cell* (2007).
151. Lang, I., Hodac, L., Friedl, T. & Feussner, I. Fatty acid profiles and their distribution patterns in microalgae: a comprehensive analysis of more than 2000 strains from the SAG culture collection. *BMC Plant Biol.* **11**, 124 (2011).
152. Kalisch, B., Dörmann, P. & Hölzl, G. DGDG and Glycolipids in Plants and Algae. *Subcell. Biochem.* **86**, 51–83 (2016).
153. Freeman Allen, C., Good, P. & Holton, R. W. Lipid Composition of Cyanidium1. *Plant Physiol.* **46**, 748–751 (1970).
154. Kleinschmidt, M. G. & McMahon, V. A. Effect of Growth Temperature on the Lipid Composition of Cyanidium caldarium: II. Glycolipid and Phospholipid Components. *Plant Physiol.* **46**, 290–293 (1970).
155. Sato, N., Mori, N., Hirashima, T. & Moriyama, T. Diverse pathways of phosphatidylcholine biosynthesis in algae as estimated by labeling studies and genomic sequence analysis. *Plant J.* **87**, 281–292 (2016).
156. Mori, N., Moriyama, T. & Sato, N. Uncommon properties of lipid biosynthesis of isolated plastids in the unicellular red alga Cyanidioschyzon merolae. *FEBS Open Bio* **9**, 114–128 (2019).
157. Sato, N., Moriyama, T., Mori, N. & Toyoshima, M. Lipid metabolism and potentials of biofuel and high added-value oil production in red algae. *World J. Microbiol. Biotechnol.* **33**, 74 (2017).
158. Kato, M., Sakai, M., Adachi, K., Ikemoto, H. & Sano, H. Distribution of betaine lipids in marine algae. *Phytochemistry* **42**, 1341–1345 (1996).
159. Iovinella, M. *et al.* Cryptic dispersal of Cyanidiophytina (Rhodophyta) in non-acidic environments from Turkey. *Extremophiles* **22**, 713–723 (2018).
160. Barcytė, D., Elster, J. & Nedbalová, L. Plastid-encoded rbc L phylogeny suggests widespread distribution of Galdieria phlegrea (Cyanidiophyceae, Rhodophyta). *Nord. J. Bot.* **36**, e01794 (2018).
161. Eren, A. *et al.* Genetic structure of Galdieria populations from Iceland. *Polar Biol.* **41**, 1681–1691 (2018).
162. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
163. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772 (2012).
164. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).

165. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
166. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
167. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
168. Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. *Bioinformatics* **24**, 1757–1764 (2008).
169. GeneiousManual.pdf.
170. Healey, A., Furtado, A., Cooper, T. & Henry, R. J. Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods* **10**, 21 (2014).
171. Maid, U. & Zetsche, K. Nucleotide sequence of the plastid 16S rRNA gene of the red alga Cyanidium caldarium. *Nucleic Acids Res.* **18**, 3996 (1990).
172. Oliveira, M. C. & Bhattacharya, D. Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of algal plastids. *Am. J. Bot.* **87**, 482–492 (2000).
173. Hou, S. *et al.* Complete genome sequence of the extremely acidophilic methanotroph isolate V4, Methylacidiphilum infernorum, a representative of the bacterial phylum Verrucomicrobia. *Biol. Direct* **3**, 26 (2008).
174. Selvaratnam, T. *et al.* Optimizing energy yields from nutrient recycling using sequential hydrothermal liquefaction with Galdieria sulphuraria. *Algal Research* **12**, 74–79 (2015).
175. Shimonaga, T. *et al.* Variation in storage alpha-glucans of the Porphyridiales (Rhodophyta). *Plant Cell Physiol.* **49**, 103–116 (2008).
176. Shimonaga, T. *et al.* Variation in storage α-polyglucans of red algae: amylose and semi-amylopectin types in Porphyridium and glycogen type in Cyanidium. *Mar. Biotechnol.* **9**, 192 (2007).
177. Stadnichuk, I. N., Semenova, L. R., Smirnova, G. P. & Usov, A. I. [A highly branched storage polyglucan in the thermoacidophilic red microalga Galdieria maxima cells]. *Prikl. Biokhim. Mikrobiol.* **43**, 88–93 (2007).